

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 12:15:47 ; Search time 103.919 Seconds
(without alignments)
1035.886 Million cell updates/sec

Title: US-10-073-293A-4
Perfect score: 1248
Sequence: 1 MESTTPQAPGSAITMEGCK.....VCGCLTALIOAFWQGPDEL 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_21:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1248	100.0	245	5 ABP57756	Abp57756 E. coli L
2	1248	100.0	245	9 ADZ39951	Adz39951 E.coli b2
3	1248	100.0	245	9 ADZ36195	Adz36195 Escherich
4	1013.5	81.2	251	7 ABO63670	Abp63670 Klebsiell
5	841.5	67.4	261	6 ABM69066	Abm69066 Photornab
6	592	47.4	266	6 ADA35238	Ada35238 Acinetoba
7	246.5	19.8	247	7 ABO61961	Abp61961 Klebsiell
8	229.5	18.4	230	5 AAG81587	Aag81587 S. epider
9	228	18.3	235	5 ABB47959	Abb47959 Listeria
10	225.5	18.1	242	5 ABP40309	Abp40309 Staphyloc
11	225.5	18.1	242	8 ADS07426	Adp07426 Staphyloc
12	219.5	17.6	230	5 ABP30530	Abp30530 Streptoco
13	219.5	17.6	230	8 ADV81752	Adv81752 Streptoco
14	219.5	17.6	231	8 ABP28562	Abp28562 Streptoco
15	219.5	17.6	231	8 ADV88332	Adv88332 Streptoco
16	219.5	17.6	231	8 ADV79585	Adv79585 Streptoco
17	214.5	17.2	255	7 ADC95293	Adc95293 E. faeciu
18	209	16.7	345	7 ABO79201	Abp79201 Pseudomon
19	204.5	16.4	238	6 ABM72815	Abm72815 Staphyloc
20	204	16.3	235	5 ABB54959	Abb54959 Lactococc
21	199.5	16.0	263	6 ADA33199	Ada33199 Acinetoba
22	197	15.8	251	4 AAB86247	Aab86247 C. glutam
23	197	15.8	251	4 AAG90035	Aag90035 C glutami
24	194.5	15.6	206	7 ADH88904	Adh88904 Enterococ

ALIGNMENTS

RESULT 1

ABP57756
ID ABP57756 standard; protein; 245 AA.
XX
AC ABP57756;
XX
DT 29-JAN-2003 (first entry)
XX
DE E. coli L-amino acid producing protein #1.
XX
KW L-amino acid; E. coli.
XX
OS Escherichia coli.
XX
PN EP1239041-A2.
XX
PD 11-SEP-2002.
XX
PF 13-FEB-2002; 2002EP-00003335.
XX
PR 13-FEB-2001; 2001RU-00103865.
PR 26-FEB-2001; 2001RU-00104998.
PR 26-FEB-2001; 2001RU-00104999.
PR 28-JUN-2001; 2001RU-00117632.
PR 28-JUN-2001; 2001RU-00117633.
XX (AJIN) AJINOMOTO CO INC.
XX Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Gussyatiner MM;
XX WPI; 2002-699856/76.
XX N-PSDB; ABV75656.
XX Novel L-amino acid producing Escherichia bacterium, is modified to
PT enhance L-amino acid production by enhancing the activities of protein
PT capable of making bacterium to have enhanced resistance to L-amino acids.
XX Disclosure; Page 23; 33pp; English.
XX The invention relates to a novel L-amino acid producing bacterium of
CC genus Escherichia, modified to enhance L-amino acid production by
CC enhancing the activities of a protein. The novel bacterium is useful for
CC producing L-amino acid e.g. L-Thr, L-Val, L-Pro, L-Met and L-Arg,
CC by cultivating the bacterium in a culture medium and collecting L-amino
CC acid to be produced and accumulated from the culture medium. The present
CC sequence represents a protein of the invention which causes increased L-
CC amino acid production in E. coli

ADB08660 Alloioioc
Adk48319 Streptoco
Aag93154 C Glutami
Abu00481 S. pneumo
Abp81300 Streptoco
Adm92060 S pneumon
ADB08658 Alloioioc
ADR94706 Novel S.
Aea58576 Streptoco
Abp66261 Bifidobac
Adf03888 Bacterial
Abo65701 Klebsiell
Ada34737 Acinetoba
Aay74752 Neisseria
Abm94911 M. xanthu
Abm72496 Staphyloc
Aay74753 Neisseria
Abm68599 Photornab
Aag91133 C Glutami
Adl65433 C. glutam
Aap70459 Sequence

25 189 15.1 226 6 ADB08660
26 187 15.0 224 8 Adk48319
27 180 14.4 237 4 AAG93154
28 178 14.3 218 6 ABU00481
29 178 14.3 218 6 ABP81300
30 178 14.3 218 6 ADM92060
31 172 13.8 214 6 ADB08658
32 171 13.7 185 8 ADR94706
33 171 13.7 185 9 AEA58576
34 167.5 13.4 338 5 ABP66261
35 146.5 11.7 230 7 ADF03888
36 143 11.5 288 7 ABO65701
37 122.5 9.8 247 6 ADA34737
38 113.5 9.1 112 3 AAY74752
39 113.5 9.1 202 9 ABM94911
40 107.5 8.6 397 6 ABM72496
41 106 8.5 113 3 AAY74753
42 105 8.4 309 6 ABM68599
43 105 8.4 324 4 AAG91133
44 105 8.4 324 7 ADL65433
45 103.5 8.3 432 1 AAP70459

XX Sequence 245 AA;
SQ
Query Match 100.0%; Score 1248; DB 5; Length 245;
Best Local Similarity 100.0%; Pred. No. 8.8e-122;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MESPTQAPGSAATFMECKDLSPIVISYIPVAFGLNATRLGFSPLSVFSCIIYAG 60
Db 1 MESPTQAPGSAATFMECKDLSPIVISYIPVAFGLNATRLGFSPLSVFSCIIYAG 60
Qy 61 ASQFVITAMLAAGSSLTAALTVMAMDVHRVLYGPSLSRSRIIQRLQSKTALMAFGLTDE 120
Db 61 ASQFVITAMLAAGSSLTAALTVMAMDVHRVLYGPSLSRSRIIQRLQSKTALMAFGLTDE 120
Qy 121 VFAAATAKLVNRNRWSENMMIGIAFSWSSWVFGTVIGAFSGSLGQGYPAVEAALGFM 180
Db 121 VFAAATAKLVNRNRWSENMMIGIAFSWSSWVFGTVIGAFSGSLGQGYPAVEAALGFM 180
Qy 191 LPALFMSFLLASFORQSLCVTAALVGLAGVTLFSPVAILAGIVCGCLTALIOAFWQ 240
Db 181 LPALFMSFLLASFORQSLCVTAALVGLAGVTLFSPVAILAGIVCGCLTALIOAFWQ 240
Qy 241 APDEL 245
Db 241 APDEL 245
RESULT 2
ADZ39951
ID ADZ39951 standard; protein; 245 AA.
XX AC ADZ39951;
XX
XX 14-JUL-2005 (first entry)
DE E.coli b2682 protein for improved amino acid production in bacteria.
XX amino acid production; fermentation; transport protein.
XX Escherichia coli.
XX EP1526179-A1.
XX 27-APR-2005.
XX
XX 13-FEB-2002; 2004EP-00028876.
XX
XX 13-FEB-2001; 2001RU-00103865.
XX 26-FEB-2001; 2001RU-00104998.
XX 26-FEB-2001; 2001RU-00104999.
XX 28-JUN-2001; 2001RU-00117632.
XX 28-JUN-2001; 2001RU-00117633.
XX 13-FEB-2002; 2002EP-00003335.
XX
XX (AJIN) AJINOMOTO CO INC.
XX
XX Tabolina EA, Rybak KV, Khoures EM, Voroshilova EB, Gusyatiner MM;
XX
XX WPI: 2005-308098/32.
XX N-PSDB; ADZ39950.
XX
XX New L-amino acid producing bacterium belonging to the genus Escherichia,
XX useful for producing L-amino acids, e.g. L-threonine, L-valine, L-
XX proline, L-methionine, or L-arginine.
XX Disclosure; SEQ ID NO 4; 35pp; English.
XX
XX The invention relates to an L-amino acid producing bacterium belonging to
XX the genus Escherichia, where the bacterium has been modified so that the
XX L-amino acid production by the bacterium is enhanced by enhancing
XX activities of proteins by transformation of the bacterium with DNA coding
XX for protein or by alteration of an expression regulation sequence of the

CC DNA on the chromosome of the bacterium. The L-amino acid producing
CC bacterium is useful for producing L-amino acids by fermentation,
CC including L-threonine, L-valine, L-proline, L-methionine, or L-arginine.
CC The genes are useful for improving L-amino acid productivity. This
CC sequence corresponds to the b2682 protein, one of the L-amino acid
CC biosynthesis pathway proteins. The protein is a putative transport
CC protein within the biosynthesis pathway.
XX
SQ Sequence 245 AA;
Query Match 100.0%; Score 1248; DB 9; Length 245;
Best Local Similarity 100.0%; Pred. No. 8.8e-122;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MESPTQAPGSAATFMECKDLSPIVISYIPVAFGLNATRLGFSPLSVFSCIIYAG 60
Db 1 MESPTQAPGSAATFMECKDLSPIVISYIPVAFGLNATRLGFSPLSVFSCIIYAG 60
Qy 61 ASQFVITAMLAAGSSLTAALTVMAMDVHRVLYGPSLSRSRIIQRLQSKTALMAFGLTDE 120
Db 61 ASQFVITAMLAAGSSLTAALTVMAMDVHRVLYGPSLSRSRIIQRLQSKTALMAFGLTDE 120
Qy 121 VFAAATAKLVNRNRWSENMMIGIAFSWSSWVFGTVIGAFSGSLGQGYPAVEAALGFM 180
Db 121 VFAAATAKLVNRNRWSENMMIGIAFSWSSWVFGTVIGAFSGSLGQGYPAVEAALGFM 180
Qy 181 LPALFMSFLLASFORQSLCVTAALVGLAGVTLFSPVAILAGIVCGCLTALIOAFWQ 240
Db 181 LPALFMSFLLASFORQSLCVTAALVGLAGVTLFSPVAILAGIVCGCLTALIOAFWQ 240
Qy 241 APDEL 245
Db 241 APDEL 245
RESULT 3
ADZ36195
ID ADZ36195 standard; protein; 245 AA.
XX AC ADZ36195;
XX
XX 14-JUL-2005 (first entry)
DE Escherichia coli b2682 protein SEQ ID NO:4.
XX amino acid production.
XX Escherichia coli.
XX EP1526181-A1.
XX 27-APR-2005.
XX
XX 13-FEB-2002; 2004EP-00028877.
XX
XX 13-FEB-2001; 2001RU-00103865.
XX 26-FEB-2001; 2001RU-00104998.
XX 26-FEB-2001; 2001RU-00104999.
XX 28-JUN-2001; 2001RU-00117632.
XX 28-JUN-2001; 2001RU-00117633.
XX 13-FEB-2002; 2002EP-00003335.
XX
XX (AJIN) AJINOMOTO CO INC.
XX
XX Tabolina EA, Rybak KV, Khoures EM, Voroshilova EB, Gusyatiner MM;
XX
XX WPI: 2005-317157/33.
XX N-PSDB; ADZ36194.
XX
XX New modified L-amino acid producing bacterium useful to enhance the
XX production of L-amino acid by enhancing the activities of proteins in a
XX cell of bacterium.

PS Disclosure; SEQ ID NO 4; 35pp; English.
 XX
 CC The invention relates to an L-amino acid (A) producing bacterium (I) (belonging to the genus Escherichia), which is modified to enhance the production of (A) by enhancing the activities of proteins (G) or (H) in a cell of (I). Also described is a method for producing (A) comprising culturing the bacterium in a culture medium and collecting the produced and accumulated L-amino acid. The modified bacterium has the ability to grow on a minimal medium containing L-amino acid or its analog in a minimal concentration and ability to grow faster on a medium containing L-amino acid or its analog than the unmodified strain or the wild type strain, or the parental strain of the bacterium. The present sequence represents the E. coli b2862 protein which has L-amino acid excretion activity.

Query Match 100.0%; Score 1248; DB 9; Length 245;
 Best Local Similarity 100.0%; Pred. No. 8.8e-122;
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 SQ Sequence 245 AA;
 QY 1 MESTPQAPGSAFMGCKDLPVIVSYIPVAFGLNATRLGFSPLSFVFFSCIIYAG 60
 DB 1 MESTPQAPGSAFMGCKDLPVIVSYIPVAFGLNATRLGFSPLSFVFFSCIIYAG 60
 QY 61 ASQFVITAMLAAGSSLMIAALTVMAMDVHRVLYGSPSLRSRIIRLQKSKTALWAFGLTDE 120
 DB 61 ASQFVITAMLAAGSSLMIAALTVMAMDVHRVLYGSPSLRSRIIRLQKSKTALWAFGLTDE 120
 QY 121 VFPAATAKLVRNRRRWSENMMIGIAFSSWSWVFTGVIAPSGSLGQYPAVEAALGFM 180
 DB 121 VFPAATAKLVRNRRRWSENMMIGIAFSSWSWVFTGVIAPSGSLGQYPAVEAALGFM 180
 QY 181 LPALFMSFLLASFORKQSLCVTAALVGLAGVTLFSPVAILAGIVCGCLTALIQAFWQ 240
 DB 181 LPALFMSFLLASFORKQSLCVTAALVGLAGVTLFSPVAILAGIVCGCLTALIQAFWQ 240
 QY 241 APDEL 245
 DB 241 APDEL 245

RESULT 4
 ABO63670
 ID ABO63670 standard; protein; 251 AA.
 AC ABO63670;
 DT 29-JUL-2004 (first entry)
 DE Klebsiella pneumoniae polypeptide seqid 10187.
 KW Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 OS Klebsiella pneumoniae.
 XX US6610836-B1.
 XX 26-AUG-2003.
 XX 27-JAN-2000; 2000US-00489039.
 XX 29-JAN-1999; 99US-0117747P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA Breton GL, Osborne M;
 PI WPI; 2003-895346/82.
 DR N-FSDB; ACH97221.
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for

PT preparing a vaccine composition against Klebsiella pneumoniae.
 PS Disclosure; SEQ ID NO 10187; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella pneumoniae polypeptide. Also described are: a recombinant expression vector comprising the nucleic acid, operably linked to a transcription regulatory element; and a cell comprising the recombinant expression vector. The nucleic acid is useful for preparing a vaccine composition against Klebsiella pneumoniae. This is the amino acid sequence of a Klebsiella pneumoniae polypeptide of the invention
 XX SQ Sequence 251 AA;

Query Match 81.2%; Score 1013.5; DB 7; Length 251;
 Best Local Similarity 81.0%; Pred. No. 3e-97;
 Matches 200; Conservative 18; Mismatches 26; Indels 3; Gaps 1;
 QY 1 MESTPQPA---PGSATFMGCKDLPVIVSYIPVAFGLNATRLGFSPLSFVFFSCII 57
 DB 4 MENPAPLTCALPERVATVGVGKDSLPIVISYLPVAFGLNATRLGFTPLSFVFFSCII 63
 QY 58 YAGASQVITAMLAAGSSLMIAALTVMAMDVHRVLYGSPSLRSRIIRLQKSKTALWAFGL 117
 DB 64 YAGASQVITAMLAAGSSLMIAALTVMAMDVHRVLYGSPSLRSRIIRLQKSKTALWAFGL 123
 QY 118 TDEVFAATAKLVRNRRRWSENMMIGIAFSSWSWVFTGVIAPSGSLGQYPAVEAAL 177
 DB 124 TDEVFAATAKLVRNRRRWSENMMIGIAFSSWSWVFTGVIAPSGSLGQYPAVEAAL 183
 QY 178 GFMLPALFMSFLLASFORKQSLCVTAALVGLAGVTLFSPVAILAGIVCGCLTALIQAF 237
 DB 184 GFMLPALFMSFLLASFORKQSLCVTAALVGLAGVTLFSPVAILAGIVCGCLTALIQAF 243
 QY 238 WQGAPE 244
 DB 244 LKGMPE 250

RESULT 5
 ABM69066
 ID ABM69066 standard; protein; 261 AA.
 AC ABM69066;
 DT 20-NOV-2003 (first entry)
 DE Photorhabdus luminescens protein sequence #2163.
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.
 OS Photorhabdus luminescens.
 XX WO200294867-A2.
 XX 28-NOV-2002.
 XX 07-FEB-2002; 2002WO-IB003040.
 XX 07-FEB-2001; 2001FR-00001659.
 XX (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PI Duchaud B, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 DR WPI; 2003-148459/14.
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,

DR WPI; 2003-895346/82.
 XX N-PSDB; ACH95512.
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 PS Disclosure; SEQ ID NO 8478; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 XX Sequence 247 AA;
 Query Match 19.8%; Score 246.5; DB 7; Length 247;
 Best Local Similarity 28.6%; Pred. No. 4.5e-17;
 Matches 67; Conservative 46; Mismatches 106; Indels 15; Gaps 6;
 QY 18 GKDSLPIVSIYIPVAFGLNATRLGSPLESVFSCIIYAGASQFVITAMLAAGSSLM 77
 Db 10 GASAIMFLCIGDFFPFIYVGLSAGSVWQSTAMSAIVTAGSAQMLNMLKTGATLG 69
 QY 78 IAAITVMMAMDVHRVLYGSPLSRSRIQRLQSKTALMAFGLTDEVFAAATAKLVRNRRWS 137
 Db 70 VLIIFTLLINLRHVLYSASI-SGTVREASPFKCFMSYALTDEVY-ATTVMKMGKKEK 127
 QY 138 ENNMIGIAFSSWSWVGTVIGAFSGSLGQYPAVEA-ALGFMLPALFMSFLLASFORK 196
 Db 128 YLFYGSAMITFWAIWVLADFLGALVGA---SFPHEIKYGLDFAMVAFAIVVP--QIK 181
 QY 197 QSLCVTAALYCALAGVTLFSP-----VAILAGIVCGCLTALIQAFWQAPDE 244
 Db 182 SQACTVAAVAAVSGVLLVLPYSLGIVASVLGVLGACVLDLAEERKQMAKTE 235
 RESULT 8
 AAG81587
 ID AAG81587 standard; protein; 230 AA.
 XX
 AC AAG81587;
 XX
 XX 03-SEP-2001 (first entry)
 XX
 XX S. epidermidis open reading frame protein sequence SEQ ID NO:268.
 DE Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
 KW endocarditis.
 XX Staphylococcus epidermidis.
 OS
 XX WO200134809-A2.
 XX
 XX 17-MAY-2001.
 PD
 XX 09-NOV-2000; 2000WO-US030782.
 PF
 XX 09-NOV-1999; 99US-0164258P.
 PR
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Kimmerly WJ;
 PI
 XX WPI; 2001-316495/33.
 DR N-PSDB; AAH52437.
 XX
 XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis.
 XX
 PS Claim 18; Page 114; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
 CC and (II) can have antibacterial activity and therefore can be used in
 CC vaccination. The nucleic acids (I) may be used to produce the S.
 CC epidermidis polypeptides (II) via the production of vectors containing
 CC them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
 CC represent oligonucleotide sequences and primers which are used in the
 CC exemplification of the present invention. N.B. The present invention
 CC specifically claims all the polynucleotide sequences given in the
 CC sequence listing of the present specification, however the sequence
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
 CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
 CC for SEQ ID NO:4455 to 4464
 XX
 XX Sequence 230 AA;
 Query Match 18.4%; Score 229.5; DB 4; Length 230;
 Best Local Similarity 28.6%; Pred. No. 2.5e-15;
 Matches 65; Conservative 47; Mismatches 102; Indels 13; Gaps 6;
 QY 14 TFMGCKDSLPIVSIYIPVAFGLNATRLGSPLESVFSCIIYAGASQFVITAMLAAG 73
 Db 5 TPKQGVKCEIPTLLTGAGVGLSGFVAVSQNFVLEITLLCLIIYAGAAQFIICTLVIAG 64
 QY 74 SSLWIAALTVMAMDVHRVLYGSPLSRSRIQRLQSKTALMAFGLTDEVFAAATAKLVRNN 133
 Db 65 TPISAIVLITLIVNSRFLSLMTLAPNYKYGFWRVGLGTL-LTDETFGVATPYVK-G 122
 QY 134 RRWSENMMIGIAFSSWSWVGTVIGAFSGSLGQYPAVEAALG--FMLPALFMSFLA 191
 Db 123 EKINDRWLHGLNITAVLFWTVSCVIGAFGE----YISNPDALGLDFAITAMFELCIS 177
 QY 192 SFQ--RKQSLCVTAALYCALAGVTLF--SIPVAILAGIVCGCLTALI 234
 Db 178 QFEGIKKSLRIYIVLVIVCVIVMMLLSLPSLYLAILIAAIVAAALL 224
 RESULT 9
 ABB47959
 ID ABB47959 standard; protein; 235 AA.
 XX
 AC ABB47959;
 XX
 XX 05-FEB-2002 (first entry)
 XX
 XX Listeria monocytogenes protein #663.
 DE
 XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 XX
 XX WO200177335-A2.
 PN
 XX 18-OCT-2001.
 PD
 XX 11-APR-2001; 2001WO-FR001118.
 PF
 XX 11-APR-2000; 2000FR-00004629.
 PR
 XX (INSP) INST PASTEUR.
 PA
 XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;

PN US2004147734-A1.
 XX 29-JUL-2004.
 XX 01-DEC-2003; 2003US-00724972.
 XX 08-NOV-1997; 97US-0064964P.
 XX 13-AUG-1998; 98US-00134001.
 XX 29-NOV-1999; 99US-00450969.
 XX (DOUC/) DOUCETTE-STAMM L.
 XX (BUSH/) BUSH D.
 XX Doucette-Stamm L, Bush D;
 XX WPI; 2004-580138/56.
 XX N-PSDB; ADS03654.
 XX New isolated polypeptide and encoding nucleic acid derived from
 PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or
 PT treating an S. epidermidis bacterial infection.
 XX Claim 17; SEQ ID NO 6721; 741pp; English.

CC The invention describes an isolated nucleic acid comprising a nucleotide
 CC sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
 CC 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any
 CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
 CC given in the specification. Also described are: a recombinant expression
 CC vector; a cell comprising a recombinant expression vector of (1);
 CC producing an S. epidermidis polypeptide; an isolated nucleic acid
 CC comprising a nucleotide sequence of at least 8 nucleotides in length; a
 CC vaccine composition for prevention or treatment of an S. epidermidis
 CC infection, comprising a nucleic acid cited above and a carrier; treating
 CC a subject for S. epidermidis infection; a recombinant or substantially
 CC pure preparation of an S. epidermidis polypeptide or its fragment; a
 CC vaccine composition for prevention or treatment of an S. epidermidis
 CC infection; detecting the presence of a Staphylococcus nucleic acid in a
 CC sample; a computer readable medium having recorded in it the nucleotide
 CC sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
 CC system for identifying fragments of the Staphylococcus genome of
 CC commercial importance; a computer based system for identifying fragments
 CC of the Staphylococcus plasmids of commercial importance; identifying
 CC commercially important nucleic acid fragments of the Staphylococcus
 CC genome and/or plasmids; and identifying an expression modulating fragment
 CC of the Staphylococcus genome and/or plasmids. The methods and
 CC compositions of the present invention are useful for the diagnosis,
 CC prevention and/or treatment of an Staphylococcus epidermidis bacterial
 CC infection. This is the amino acid sequence of a S. epidermidis protein of
 CC the invention.

XX Sequence 242 AA;

Query Match 18.1%; Score 225.5; DB 8; Length 242;
 Best Local Similarity 28.6%; Pred. No. 6.9e-15;
 Matches 66; Conservative 39; Mismatches 89; Indels 37; Gaps 7;
 QY 14 TMEGCKDSLPIVTSYIPVAFAGLNATRLGFSPLSFVFFSCIIYAGASQFVITAMLAAG 73
 Db 17 TFKQGVKECTPLIGYAGVGLSPGIVAAASQNFVLEIILLCIIYAGAAQFICTLVIA 76
 QY 74 SSLWIAALTVMADMVRHVLVGPSSLRRIIQRLOKSKTALWAFGLTDEVFMAATAKLVN 133
 Db 77 TPISAIVLTLLVNSRMFLLSMTLAPNKKYQGFNRVGLTGLT-LTDETFGVAITPYVK-G 134
 QY 134 RRNSENMIGIAFSSWSWVFGTVIGAFSGSLGQYPAVEAIG--FMLPALFMSFLA 191
 Db 135 EKINDRWLHGLNITAYLFTWVSCVIGAIFGE-----YISNPDALGLDFAITAMFIFLCIS 189
 QY 192 SFQ--RKQSL-----CV-----TAAVLCALAGVTL 214
 Db 190 QFEGIKRSRRIYIIVLIVCVIVMMLLSLILSPSVAILIAIIVAAALGVVM 240

RESULT 12
 ABP30530
 ID ABP30530 standard; protein; 230 AA.
 XX AC ABP30530;
 XX 02-JUL-2002 (first entry)
 XX Streptococcus polypeptide SEQ ID NO 10236.
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX OS Streptococcus agalactiae.
 XX WO200234771-A2.
 XX 02-MAY-2002.
 XX 29-OCT-2001; 2001WO-GB004789.
 XX 27-OCT-2000; 2000GB-00026333.
 XX 24-NOV-2000; 2000GB-00028727.
 XX 07-MAR-2001; 2001GB-00005640.
 XX (CHIR-) CHIRON SPA.
 XX (GENO-) INST GENOMIC RES.
 XX Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;
 PI Tettelin H;
 XX WPI; 2002-352536/38.
 XX N-PSDB; ABN71161.
 XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX Claim 1; Page 4153; 4525pp; English.
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX Sequence 230 AA;
 Query Match 17.6%; Score 219.5; DB 5; Length 230;
 Best Local Similarity 29.8%; Pred. No. 2.8e-14;
 Matches 67; Conservative 38; Mismatches 87; Indels 39; Gaps 7;
 QY 15 FMEGCKDSLPIVTSYIPVAFAGLNATRLGFSPLSFVFFSCIIYAGASQFVITAMLAAGS 74
 Db 5 FKEGVKDALPTALGYISIGLAFGIVASASDLSAIEVGLMSALVYGSAQFAMCALLAKA 64
 QY 75 SLWIAALTVMADMVRHVLVGPSSLRRIIQRLOKSKTALWAFG--LTDEVFMAATAKLVN 132
 Db 65 DLMTITMTVFLNLRNLM--SLHATTIFKSAHLMNQL-AIGTLITDESYGVLIGAL-H 120

QY 133 NRRSENWMIIGIARSSSWVFGTVIGAFSGSLGLOQYPAVEA-ALGFMLPALFMS----- 187
 Db 121 HKVWSPSMHGNVMSYLTWISTIIIGTLGSTI-----PNPEMFGDLPALVAMFGLPVF 176
 QY 188 -----FLASFKQKSLCVTAALVGLAGVTL 214
 Db 177 QLFQMLSDGKRLVYVVLASVGLSYFLLATFLSGALSVLLATVVGCSGVVL 227

RESULT 13
 ID ADV81752 standard; protein; 230 AA.
 XX ADV81752;
 AC
 XX
 DT 24-FEB-2005 (first entry)
 XX
 DE Streptococcus agalactiae protein, SEQ ID 2893.
 XX
 KW Antibacterial; vaccine; bacterial infection.
 XX
 OS Streptococcus agalactiae.
 XX
 XX WO200292818-A2.
 PN
 XX
 XX 21-NOV-2002.
 PD
 XX
 PF 26-APR-2002; 2002WO-IB003059.
 XX
 XX 26-APR-2001; 2001FR-00005642.
 PR
 XX
 XX (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Glaeser P, Rueniok C, Chevalier F, Frangeul L, Lalioui L;
 PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Cuot P, Kunst F;
 PI WPI; 2004-101891/11.
 DR
 XX
 XX Genomic nucleotide sequences encoding polypeptides of Streptococcus
 PT agalactiae for the development of vaccines, diagnostic tools, DNA chips
 PT and identification of therapeutic targets.
 XX
 XX Claim 6; SEQ ID NO 2893; 439pp; French.
 PS
 XX
 XX The present invention relates to novel Streptococcus agalactiae
 CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and
 CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The
 CC nucleotide sequences encode polypeptides of S. agalactiae involved in the
 CC synthesis of amino acids, cell membranes, intermediate (central)
 CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,
 CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,
 CC regulatory functions, replication, transcription, translation, protein
 CC transport, adaptation to atypical conditions, sensitivity to medicines
 CC and/or analogues, functions related to transposons, biosynthesis of
 CC cofactors, prosthetic groups and transporters, cell membrane proteins and
 CC cellular machinery. (I) are useful for the detection and/or amplification
 CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are
 CC useful for treatment of a bacterial S. agalactiae infection. The complete
 CC genome of Streptococcus agalactiae is given in ADV81204. Note: The
 CC present patent is an equivalent for the basic patent FR2824074A1, which
 CC contains only 2344 sequences.
 XX
 SQ Sequence 230 AA;

Query Match 17.6%; Score 219.5; DB 8; Length 230;
 Best Local Similarity 29.0%; Pred. No. 2.8e-14;
 Matches 67; Conservative 38; Mismatches 87; Indels 39; Gaps 7;

QY 15 FMGCKDSLPIVISYIPAFAGFLNATRLGFLSPLESVFFSCIIYAGASQFVTMLAAGS 74
 Db 5 FKGVKDALPTALGVISIGLAFGIVASASDLSAIEVGLMSALVYGSAQFAMCALLAKA 64

QY 75 SLWIAALTVMAMDRVHLYGSLRSRIQRLQKSKTALWAFG--LTDEVFPAATAKLVRN 132
 Db 65 DLMTITVTVFLVNLRNMLM--SLHATTIFKSAHLMNQL-AIGTLITDESYGVLGEAL-H 120
 QY 133 NRRSENWMIIGIARSSSWVFGTVIGAFSGSLGLOQYPAVEA-ALGFMLPALFMS----- 187
 Db 121 HKVWSPSMHGNVMSYLTWISTIIIGTLGSTI-----PNPEMFGDLPALVAMFGLPVF 176
 QY 188 -----FLASFKQKSLCVTAALVGLAGVTL 214
 Db 177 QLFQMLSDGKRLVYVVLASVGLSYFLLATFLSGALSVLLATVVGCSGVVL 227

RESULT 14
 ABP28562
 ID ABP28562 standard; protein; 231 AA.
 XX
 AC ABP28562;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 6300.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 XX WO200234771-A2.
 PN
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB004789.
 XX
 PR 27-OCT-2000; 2000GB-00026333.
 XX
 PR 24-NOV-2000; 2000GB-00028727.
 PR
 PR 07-MAR-2001; 2001GB-00005640.
 XX
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 XX Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;
 PI Tettelin H;
 PI
 XX
 XX WPI; 2002-352536/38.
 DR
 DR N-PSDB; ABN69193.
 XX
 XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX
 XX PS
 XX
 XX Claim 1; Page 3795; 4525pp; English.
 CC
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX

```

CC equivalent for the present basic patent FR2824074A1. WO200292818A2
CC contains 6617 sequence whereas the present patent only contains 2344
CC sequences.
XX
SQ Sequence 231 AA;

Query Match      17.6%; Score 219.5; DB 8; Length 231;
Best Local Similarity 29.0%; Pred. No. 2.8e-14;
Matches 67; Conservative 38; Mismatches 87; Indels 39; Gaps 7;

QY 15 FMEGCKDSLPIVISYIPVAFAPAGLNATRLGFPSPLESVPFSCIIYAGASOFVITAMLAGS 74
Db 6 FREGVKDALPTALGYISIGLAFGIVASASDLSAIEVGLMSALVYGSAOFAMCALLAKA 65

QY 75 SLWIAALTYMAMDVHRHLYGSPSLRSRIIQRLQSKKTALWAFG--LTDEVFAAATAKLVRN 132
Db 66 DLMTIITWTVLVNLRNLM--SLHATTIFKSAHLMNQL-AIGTLITDES YGVLLGEAL-H 121

QY 133 NRRWSENWMIIGIAFPSSWSNWVPGTVIGAFSGGLLQGYPAVEA-ALGFMLPALFMS---- 187
Db 122 HKVVPSPWMHGNNVMSYLTWVISTIIIGTLIGSTI----PNPEMFGLDLALVAMFIFGLFVF 177

QY 188 -----FLASFORQSLCVTAALYGALAGVTL 214
Db 178 QLFGMLSDGKRLVVVYVVASVGLSYFLLATFLSGALSVLLATVVGCSVGWVL 228

Search completed: February 15, 2006, 12:18:31
Job time : 106.919 secs

```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 12:18:47 ; Search time 33.0337 Seconds
(without alignments)
713.608 Million cell updates/sec

Title: US-10-073-293A-4
Perfect score: 1248
Sequence: 1 MESPTPOPAPGSATFMEGCK.....VCGCLTALIOAFWQGAPEL 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1248	100.0	245	1 C65048	hypothetical prote
2	1247	99.9	245	2 C85916	hypothetical prote
3	1247	99.9	245	2 H91071	hypothetical prote
4	871.5	69.8	257	2 AF0396	probable amino aci
5	316.5	25.4	219	1 B69469	conserved hypotet
6	287	23.0	237	2 F84013	branched-chain ami
7	232	18.6	240	2 B97608	azlc family protei
8	232	18.6	240	2 AE2830	conserved hypotet
9	228	18.3	235	2 AB1255	transport proteins
10	225	18.0	235	2 AG1617	transport protein
11	212.5	17.0	235	2 G75494	Azlc family protei
12	209	16.7	252	2 H83390	hypothetical prote
13	208	16.7	239	2 G82388	Azlc family protei
14	204	16.3	235	2 E86826	amino acid permea
15	203.5	16.3	231	2 B89759	hypothetical prote
16	201	16.1	234	2 B97492	azlc family protei
17	201	16.1	234	2 AI2709	branched-chain ami
18	195.5	15.7	228	2 F71831	hypothetical prote
19	191.5	15.3	228	1 C64686	conserved hypotet
20	178	14.3	218	2 G95016	conserved hypotet
21	168	13.5	254	1 G69592	branched-chain ami
22	153	12.3	219	2 AE0945	probable exported
23	146	11.7	246	2 AD3574	branched-chain ami
24	145.5	11.7	256	2 T35133	hypothetical prote
25	136	10.9	244	1 F64041	hypothetical prote
26	135	10.8	275	2 D97525	hypothetical prote
27	124.5	10.0	160	2 C81877	hypothetical prote
28	109	8.7	1108	2 A48508	cyclic-nucleotide
29	107.5	8.6	397	2 G90048	hypothetical prote

30	104	8.3	651	2	S44257	phosphotransferase
31	103.5	8.3	300	2	E87280	rard protein [impo
32	103.5	8.3	410	2	AH0051	probable O-antigen
33	103.5	8.3	432	2	S67821	gumE protein - Xan
34	102	8.2	343	2	B69800	iron(III) dicitrat
35	101	8.1	467	2	E95850	probable amino aci
36	101	8.1	583	2	T12576	probable phosphate
37	100	8.0	406	2	A71109	probable phosphate
38	100	8.0	512	2	B70786	branched-integral
39	99.5	8.0	307	2	B36125	branched-chain ami
40	99	7.9	463	2	C69751	sodium/proton-depe
41	98.5	7.9	465	2	I39473	Na+-dependent phos
42	98.5	7.9	467	2	A48916	sodium phosphate t
43	98	7.9	429	2	B83826	hypothetical prote
44	97.5	7.8	460	2	A84154	amino acid transpo
45	97.5	7.8	517	2	AC2070	Na+/H+ antiporter

ALIGNMENTS

RESULT 1

C65048
hypothetical protein b2682 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C;Accession: C65048
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID: 97426617; PMID: 9278503
A;Accession: C65048
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-245 <BLAT>
A;Cross-references: UNIPROT:P76630; UNIPARC:UPI000004F5DD; GB:AE000353; GB:U00096; NID
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: branched-chain amino acid transport protein, Azlc type

Query Match 100.0%; Score 1248; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 3.8e-96;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MESPTPOPAPGSATFMEGCKDSLPIVTSYIPVAFGLNATRLGFSPLSFVFFSIIYAG	60
Db	1	MESPTPOPAPGSATFMEGCKDSLPIVTSYIPVAFGLNATRLGFSPLSFVFFSIIYAG	60
Qy	61	ASQFVITAMLAAGSSLMWIAALTVMAMDVRHVLVYGPSLRSRIIQRLQSKTALWAFGLTDE	120
Db	61	ASQFVITAMLAAGSSLMWIAALTVMAMDVRHVLVYGPSLRSRIIQRLQSKTALWAFGLTDE	120
Qy	121	VFAAATAKLVRNRRNSNNMIGIAFSSWSWVFGTVIGAFSGSLGQGPVFAAALGFM	180
Db	121	VFAAATAKLVRNRRNSNNMIGIAFSSWSWVFGTVIGAFSGSLGQGPVFAAALGFM	180
Qy	181	LPALFWSFLASFORQSLCVTAALVAGLVTLFSTIPVAILAGIVCGCLTALIOAFWQ	240
Db	181	LPALFWSFLASFORQSLCVTAALVAGLVTLFSTIPVAILAGIVCGCLTALIOAFWQ	240
Qy	241	APDEL 245	
Db	241	APDEL 245	

RESULT 2

C85916
hypothetical protein Z3983 [imported] - Escherichia coli (strain O157:H7, substrain EDL
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 05-Oct-2004
C;Accession: C85916
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca

Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85916
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-245 <STO>
A:Cross-references: UNIPROT:C8X908; UNIPROT:Q8FEQ6; UNIPARC:UPI00000D0899; GB:AE005174;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z3983
C:Superfamily: branched-chain amino acid transport protein, AzlC type

Query Match 99.9%; Score 1247; DB 2; Length 245;
Best Local Similarity 99.6%; Pred. No. 4.6e-96;
Matches 244; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESPTQPAPGSATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAG 60
Db 1 MESPTQPAPGSATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAG 60

Qy 61 ASQFVTAMLAAGSSLIWTAALTVAMDVHRVLYGFSLSRSRIIQRLOKSKTALWAFGLTDE 120
Db 61 ASQFVTAMLAAGSSLIWTAALTVAMDVHRVLYGFSLSRSRIIQRLOKSKTALWAFGLTDE 120

Qy 121 VFAAATAKLVRRNRWSENWMIIGIAFSSWSWVFGTVIGAFSGSLLOQGYPAVEAALGFM 180
Db 121 VFAAATAKLVRRNRWSENWMIIGIAFSSWSWVFGTVIGAFSGSLLOQGYPAVEAALGFM 180

Qy 181 LPALFMSFLLASFORKQSLCVTAALVAGALAGVTLPSPVAILAGIVCGCLTALIQAFWQ 240
Db 181 LPALFMSFLLASFORKQSLCVTAALVAGALAGVTLPSPVAILAGIVCGCLTALIQAFWQ 240

Qy 241 APDEL 245
Db 241 APDEL 245

RESULT 3
H91071
hypothetical protein ECs3544 [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 05-Oct-2004
C:Accession: H91071
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference strain O157:H7, substrain R1
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H91071
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-245 <HAY>
A:Cross-references: UNIPROT:C8X908; UNIPROT:Q8FEQ6; UNIPARC:UPI00000D0899; GB:BA000007;
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs3544
C:Superfamily: branched-chain amino acid transport protein, AzlC type

Query Match 99.9%; Score 1247; DB 2; Length 245;
Best Local Similarity 99.6%; Pred. No. 4.6e-96;
Matches 244; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESPTQPAPGSATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAG 60
Db 1 MESPTQPAPGSATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAG 60

Qy 61 ASQFVTAMLAAGSSLIWTAALTVAMDVHRVLYGFSLSRSRIIQRLOKSKTALWAFGLTDE 120
Db 61 ASQFVTAMLAAGSSLIWTAALTVAMDVHRVLYGFSLSRSRIIQRLOKSKTALWAFGLTDE 120

Qy 121 VFAAATAKLVRRNRWSENWMIIGIAFSSWSWVFGTVIGAFSGSLLOQGYPAVEAALGFM 180
Db 121 VFAAATAKLVRRNRWSENWMIIGIAFSSWSWVFGTVIGAFSGSLLOQGYPAVEAALGFM 180

Db 121 VFAAATAKLVRRNRWSENWMIIGIAFSSWSWVFGTVIGAFSGSLLOQGYPAVEAALGFM 180

Qy 181 LPALFMSFLLASFORKQSLCVTAALVAGALAGVTLPSPVAILAGIVCGCLTALIQAFWQ 240
Db 181 LPALFMSFLLASFORKQSLCVTAALVAGALAGVTLPSPVAILAGIVCGCLTALIQAFWQ 240

Qy 241 APDEL 245
Db 241 APDEL 245

RESULT 4
AF0396
probable amino acid transporter YPO3264 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 05-Oct-2004
C:Accession: AF0396
R:Parikhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0396
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <KUR>
A:Cross-references: UNIPROT:Q8ZBX0; UNIPARC:UPI00000DCM45; GB:AL590842; PIDN:CNC92498.1;
C:Genetics:
A:Gene: YPO3264
C:Superfamily: branched-chain amino acid transport protein, AzlC type

Query Match 69.8%; Score 871.5; DB 2; Length 257;
Best Local Similarity 70.9%; Pred. No. 5.9e-65;
Matches 168; Conservative 29; Mismatches 35; Indels 5; Gaps 1;

Qy 4 PTPQAPGS-----ATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIY 58
Db 2 PNPTAFSSSTSAATFVEGITDSLPIVIGYLPVAFAGLSAVKLGFTPLSFVFFSCIIY 61

Qy 59 AGASQFVITAMLAAGSSLIWTAALTVAMDVHRVLYGFSLSRSRIIQRLOKSKTALWAFGLT 118
Db 62 AGASQFVITAMLAAGSSLIWTAALTVAMDVHRVLYGFSLSRSRIIQRLOKSKTALWAFGLT 121

Qy 119 DEVFAAATAKLVRRNRWSENWMIIGIAFSSWSWVFGTVIGAFSGSLLOQGYPAVEAALG 178
Db 122 DEVFAAATAKLVRRNRWSENWMIIGIAFSSWSWVFGTVIGAFSGSLLOQGYPAVEAALG 181

Qy 179 FMLPALFMSFLLASFORKQSLCVTAALVAGALAGVTLPSPVAILAGIVCGCLTALIQ 235
Db 182 FMLPALFMSFLLASFORKQSLCVTAALVAGALAGVTLPSPVAILAGIVCGCLTALIQ 238

RESULT 5
B69469
conserved hypothetical protein AF1755 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C:Accession: B69469
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: B69469
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-219 <KLE>
A:Cross-references: UNIPROT:O28519; UNIPARC:UPI0000056BED; GB:AE000982; NIT
C:Superfamily: branched-chain amino acid transport protein, AzlC type

Query Match 25.4%; Score 316.5; DB 1; Length 219;
Best Local Similarity 35.8%; Pred. No. 4.4e-19;
Matches 82; Conservative 44; Mismatches 86; Indels 17; Gaps 6;

QY 12 SATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAGASQFVITAMLA 71
DB 7 SAMFRKGLVYFPPIVMAIYIPVAFVGLARTLGFSEVMALSLIIFAGASQFALITLYS 66

QY 72 AGSSLWIAALTVMAMDVHRVLYGFSLSRIIQRLOKSKTALMAFGLTDEVFAAATAKLVR 131
DB 67 --QSLLSAIPIPIFLNLRHIYS---SIIAQLKLRPHISAFGLTDEVFAVSV----- 115

QY 132 NNRWSENWMIATFSSWSWVFTVIGAFSGGLQGYPAVEAALGFMPLPALFMSFLLA 191
DB 116 -NSAENRFLGLGSLYSWVGGTALGVLAGSTLILDRD-VYSGALFVPSISALEFLVLPL 173

QY 192 SFORKQSLCVTAALVGLAGVTLSPVAILAGIVCGCLTALIOAFWOG 240
DB 174 NLKGRH---VRAAVSGGVALA-FHLLNLTSVGIIAALAGPLLSGWDG 218

RESULT 6
F84013
branched-chain amino acid transporter BH2910 [imported] - Bacillus halodurans (strain C-
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 05-Oct-2004
C/Accession: F84013
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: F84013
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-237 <STO>
A/Cross-references: UNIPROT:Q9K8U2; UNIPARC:UPI00000C0406; GB:AP001517; GB:BA000004; NID
A/Experimental source: strain C-125
C/Genetics:
A/Supergene: BH2910
C/Supergene: branched-chain amino acid transport protein, AzLC type

Query Match 23.0%; Score 287; DB 2; Length 237;
Best Local Similarity 32.1%; Pred. No. 1.3e-16;
Matches 70; Conservative 49; Mismatches 87; Indels 12; Gaps 5;

QY 1 MESPTQPAGSATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAG 60
DB 1 METASNMKARETDWLRGAIGSIAIGYIPASITFGLLAGTGLTLETTIAMSLFVYAG 60

QY 61 ASQFVITAMLAAGSSLMIAALTVMAMDVHRVLYGFSLSRIIQRLOKSKTALMAFGLTDE 120
DB 61 ASQYMALTLIAIGTGTEIILTTIVINRHLMSASIRAK-MEDTHPVKRIATAFGLTDE 119

QY 121 VFAAATAKLVRNRRWSENWMIATFSSWSWVFTVIGAFSGGLQGYPA-VEAALGF 179
DB 120 VFALVTSQ---DRRLTNGFVIGAVIATVSVVHSAVGYIVGSAL----PATLQQGMV 171

QY 180 MLPALFMSFLLASFORQSLCVTAALVGLAGVTLSPVAILAGIVCGCLTALIOAFWOG 217
DB 172 ALYAMFALLIPSRVKKRHSVLILAGTAALLNG--LFSL 207

RESULT 7
B97608
azlc family protein [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 05-Oct-2004
C/Accession: B97608
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: B97608
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-240 <KUR>
A/Cross-references: UNIPROT:Q8UDQ3; UNIPARC:UPI00000D1D87; GB:AE007869; PIDN:AAK87819.1
C/Genetics:
A/Supergene: AGR_C_3745
C/Supergene: branched-chain amino acid transport protein, AzLC type

Query Match 18.6%; Score 232; DB 2; Length 240;
Best Local Similarity 30.2%; Pred. No. 4.8e-12;
Matches 68; Conservative 37; Mismatches 112; Indels 8; Gaps 6;

QY 12 SATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAGASQFVITAMLA 71
DB 3 NADFREGKSGFPFIALSAAPFGALFGAVAVDNGLSITEATIMSGTVYAGASQVGLF 62

QY 72 AGSSLWIAALTVMAMDVHRVLYGFSLSRIIQRLOKSKTALMAFGLTDEVFAAATAKLVR 131
DB 63 QKVAPMLVLSVFAVNFRLHLYSAAI-ARMISNWSLQKAAGFFVLVDPQFAESVRK-YE 120

QY 132 NNRWSENWMIATFSSWSWVFTVIGAFSGGLQGYPAVEAALGFMPLPALFMSFLLA 191
DB 121 NTGTGVGFSWMTGFPATPVYVILWMLAMTILGASLGN--LVGDPKA-IGLDVLLPIYFMGMVL- 176

QY 192 SFORKQSL--CVTAALVGLAGVTLSPVAILAGIVCGCLTALI 234
DB 177 SFQRNFYFVPLASAGATVAFVGVSPHVSILGAVAGIVAVL 221

RESULT 8
AE2830
conserved hypothetical protein Atu2068 [imported] - Agrobacterium tumefaciens (strain C-
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 05-Oct-2004
C/Accession: AE2830
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AE2830
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-240 <KUR>
A/Cross-references: UNIPROT:Q8UDQ3; UNIPARC:UPI00000D1D87; GB:AE008688; PIDN:AAL43059.1
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Supergene: Atu2068
C/Supergene: branched-chain amino acid transport protein, AzLC type

Query Match 18.6%; Score 232; DB 2; Length 240;
Best Local Similarity 30.2%; Pred. No. 4.8e-12;
Matches 68; Conservative 37; Mismatches 112; Indels 8; Gaps 6;

QY 12 SATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAGASQFVITAMLA 71
DB 3 NADFREGKSGFPFIALSAAPFGALFGAVAVDNGLSITEATIMSGTVYAGASQVGLF 62

QY 72 AGSSLWIAALTVMAMDVHRVLYGFSLSRIIQRLOKSKTALMAFGLTDEVFAAATAKLVR 131
DB 63 QKVAPMLVLSVFAVNFRLHLYSAAI-ARMISNWSLQKAAGFFVLVDPQFAESVRK-YE 120

QY 132 NNRWSENWMIATFSSWSWVFTVIGAFSGGLQGYPAVEAALGFMPLPALFMSFLLA 191
DB 121 NTGTGVGFSWMTGFPATPVYVILWMLAMTILGASLGN--LVGDPKA-IGLDVLLPIYFMGMVL- 176

Qy 62 SQFVITAMLAAGSSLWIAALVTWAMDVHRVLYGSPLSRRIQRLQKSKTALWAFGLTDEV 121
Db 65 AQLVAIGMFTKGAGLLSLLIATFTTIGRHFLYSVSMRSKISPLPLURWRLTL--GFLLTDBEL 123

Qy 122 FAAATAKLVRNRRWSENMWIGIAFSSWSSWVFTGIVAFSGSGLQGYPAV-EAALGPM 180
Db 124 FATCGASDQGFNRW---YALGAGLSFYLINLASLVGIVAGSYL----PDLNQNGLEPA 176

Qy 181 LPAIFMSFLLASFORKOSL--CVTAAALVGLA-----AGVTLSFIPVAAILAG 224
Db 177 VAATFIAIVIPNIKSWPVLISVLTALVLSVLTVMGIEGSLMFASIGAMLAG 228

RESULT 14
E86826
amino acid permease yqfd [imported] - Lactococcus lactis subsp. lactis (strain C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-May-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: E86826
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, T.
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: E86826
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-235 <STO>
A;Cross-references: UNIPROT:Q9CF68; UNIPARC:UPI00000C6A96; GB:AE005176; PID:g1
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yqfd

Query Match 16.3%; Score 204; DB 2; Length 235;
Best Local Similarity 29.4%; Pred. No. 9.9e-10;
Matches 59; Conservative 35; Mismatches 101; Indels 6; Gaps 3

Qy 14 TMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSEVFSCIIYAGASQFVITAMLAAG 73
Db 6 TFKQGLKDNPTVFGYIGIGIAGFMGHSGFSVWVILLSLIVAGSAQFIMVSLMATH 65

Qy 74 SSLWIAALVTWAMDVHRVLYGSPLSRRIQRLQKSKTALWAFGLTDEVFAAATAKLVRNN 133
Db 66 SPIMSVLSVFLVNSRIILMSMTTAS-YFKNESLLKNLLGTLTDESFGALGMNKQNYTE 124

Qy 134 RWSENMWIGIAFSSWSSWVFTGIVAFSGSGLQGYPAV-EAALGPMALPMPFMSPLAS 193
Db 125 GKLNFSWFNASNLAYLVWALASAGLGNLLAN---PEKGLGFGFVIANFILLGLYQL 181

Qy 194 QRKQSLCVTAAALVGLAGVTL 214
Db 182 ISDKTGLMLQLV--WVGITL 200

RESULT 15
B89759
hypothetical protein SA0010 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: B89759
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Seki
C.; Shiba, T.; Hattori, M.; Ogaewara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89759; MUID:21311952; PMID:11418146
A;Accession: B89759
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-231 <KUR>
A;Cross-references: UNIPROT:Q99XG1; UNIPARC:UPI00000C3AA13; GB:BA000018; PID:g1
A;Experimental source: strain N315
C;Genetics:

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 12:16:07 ; Search time 203.02 Seconds
(without alignments)
851.417 Million cell updates/sec

Title: US-10-073-293A-4
Perfect score: 1248
Sequence: 1 MESPTQPAPGSAFMFGCK.....VCGCLTALIOAFMQGAPDEL 245

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1248	100.0	245	1 YGAZ_ECOLI	P76630 escherichia
2	1247	99.9	245	2 Q8FEQ6_ECOLI	Q8FEQ6 escherichia
3	1247	99.9	245	2 Q8X908_ECO57	Q8X908 escherichia
4	1243	99.6	245	2 Q83J26_SHIFL	Q83J26 shigella fl
5	871.5	69.8	257	2 Q8ZBX0_YERPE	Q8ZBX0 yersinia pe
6	870.5	69.8	257	2 Q66E34_YERFS	Q66E34 yersinia fs
7	845.5	67.7	243	2 Q6DID9_ERWCT	Q6DID9 erwinia car
8	841.5	67.4	252	2 Q7N779_PHOLL	Q7N779 photorhabdu
9	834	66.8	247	2 Q93KB4_ERWCH	Q93KB4 erwinia chr
10	385	30.8	237	2 Q7WB32_BORPA	Q7WB32 bordetella
11	385	30.8	237	2 Q7WMJ8_BORBR	Q7WMJ8 bordetella
12	316.5	25.4	219	1 Y1755_ARCFU	Y1755 archaeoglob
13	308.5	24.7	224	2 Q8VW59_BRUBA	Q8VW59 brucella ab
14	308.5	24.7	224	2 Q8FY15_BRUSU	Q8FY15 brucella su
15	301.5	24.2	238	2 Q6LZP2_METWP	Q6LZP2 methanococ
16	287	23.0	237	2 Q9K8U2_BACHD	Q9K8U2 bacillus ha
17	286.5	23.0	216	2 Q5SLQ8_THET8	Q5SLQ8 thermus the
18	286.5	23.0	216	2 Q72GU7_THET2	Q72GU7 thermus the
19	281	22.5	238	2 Q5L272_GEOKA	Q5L272 geobacillus
20	276	22.1	234	2 Q67SA1_SYMTB	Q67SA1 symbiobacte
21	269	21.6	241	2 Q6HKN2_BACHK	Q6HKN2 bacillus th
22	269	21.6	241	2 Q81SD8_BACAN	Q81SD8 bacillus an
23	268.5	21.5	238	2 Q65LQ0_BACLD	Q65LQ0 bacillus li
24	268	21.5	241	2 Q63D72_BACCC	Q63D72 bacillus ce
25	267	21.4	241	2 Q4MUB4_BACCE	Q4MUB4 bacillus ce
26	266	21.3	241	2 Q73A13_BACCI	Q73A13 bacillus ce
27	262	21.0	241	2 Q81FC0_BACCR	Q81FC0 bacillus ce
28	254	20.4	242	2 Q731S3_BACCI	Q731S3 bacillus ce
29	251	20.1	233	2 Q6AKD1_DESFS	Q6AKD1 desulfocale
30	248	19.9	243	2 Q5WAZ4_BACSK	Q5WAZ4 bacillus cl
31	247.5	19.8	242	2 Q8Y223_RALSO	Q8Y223 ralstonia s

32	243	19.5	243	2	Q4FS78_9GAMM	Q4fs78 psychrobact
33	240.5	19.3	235	2	Q831J7_ENTFA	Q831j7 enterococcu
34	238	19.1	234	2	Q74F41_GEOSL	Q74f41 geobacter s
35	236	18.9	242	2	Q98KC6_RHILO	Q98kc6 rhizobium l
36	236	18.9	243	2	Q4NSK9_9DELT	Q4nsk9 anaeromyxob
37	235.5	18.9	230	2	Q8EG49_SHEON	Q8eg49 shewanella
38	235	18.8	232	2	Q4FL99_9RICK	Q4fl99 candidatus
39	234	18.8	232	2	Q5LY33_STRT1	Q5ly33 streptococc
40	234	18.8	232	2	Q5M2N8_STRT2	Q5m2n8 streptococc
41	232	18.6	240	2	Q8UDQ3_AGRTS	Q8udq3 agrobacteri
42	231.5	18.5	232	2	Q8DVS9_STRMU	Q8dvs9 streptococc
43	228.5	18.3	241	2	Q8CUI8_OCEIH	Q8cul8 oceanobacil
44	228	18.3	230	2	Q5V111_HALMA	Q5v111 haloarcula
45	228	18.3	235	2	Q8Y761_LISMO	Q8y761 listeria mo

ALIGNMENTS

RESULT 1

ID	YGAZ_ECOLI	STANDARD	PRT	245 AA
AC	P76630			
DT	15-JUL-1998	(Rel. 36, Created)		
DT	15-JUL-1998	(Rel. 36, Last sequence update)		
DT	13-SEP-2005	(Rel. 48, Last annotation update)		
DE	Inner membrane protein ygaZ.			
GN	.Name=ygaZ; OrderedLocustNames=b2682;			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
NC	NCBI_TaxID=562;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=K12 / MG1655;			
RC	MDLINF=9746617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.";			
RL	Science 277:1453-1474 (1997).			
RN	[2]			
RP	TOPOLOGY.			
RC	STRAIN=K12 / MG1655;			
RC	PubMed=15919996; DOI=10.1126/science.1109730;			
RA	Daley D.O., Rapp M., Graneth E., Melen K., Drew D., von Heijne G.;			
RT	"Global topology analysis of the Escherichia coli inner membrane proteome.";			
RL	Science 308:1321-1323(2005).			
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.			
CC	-I- SIMILARITY: Belongs to the azlC family.			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.			
CC	EMBL; U00096; AAC75729.1; -; Genomic_DNA.			
DR	PIR; C65048; C65048.			
DR	EcoBASE; EB3299; -.			
DR	EcoGene; EG13528; YgaZ.			
DR	InterPro; IPR011606; AzlC-like.			
DR	InterPro; IPR012294; TFIID_C/glycos_N.			
DR	Pfam; PF03591; AzlC; 1.			
DR	Complete proteome; Inner membrane; Membrane; Transmembrane; Transport.			
FT	TOPO_DOM 1 24 Cytoplasmic (Potential).			
FT	TRANSMEM 25 45 Potential.			
FT	TOPO_DOM 46 63 Periplasmic (Potential).			
FT	TRANSMEM 64 84 Potential.			
FT	TOPO_DOM 85 109 Cytoplasmic (Potential).			
FT	TRANSMEM 110 130 Potential.			

```

FT TOPO_DOM 131 140 Periplasmic (Potential).
FT TRANSMEM 141 161 Potential.
FT TOPO_DOM 162 172 Cytoplasmic (Potential).
FT TRANSMEM 173 193 Potential.
FT TOPO_DOM 194 205 Periplasmic (Potential).
FT TRANSMEM 206 226 Potential.
FT TOPO_DOM 227 245 Cytoplasmic (Potential).
SQ SEQUENCE 245 AA; 26108 MW; 22ACB8AB8D7D651B CRC64;

Query Match 100.0%; Score 1248; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 7.7e-93;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESPTQPAPGSAATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIYAG 60
Db 1 MESPTQPAPGSAATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIYAG 60

Qy 61 ASQFVITAMLAAGSSLWIAALTVMAMDVRHLYGFSLSRSRIIQRLQSKTALWAFGLTDE 120
Db 61 ASQFVITAMLAAGSSLWIAALTVMAMDVRHLYGFSLSRSRIIQRLQSKTALWAFGLTDE 120

Qy 121 VFAAATAKLVRRNRRWSNMMGIAFSSWSWVFTVIGAFSGSLGQGYPAVEAALGFM 180
Db 121 VFAAATAKLVRRNRRWSNMMGIAFSSWSWVFTVIGAFSGSLGQGYPAVEAALGFM 180

Qy 181 LPALFMSFLLASFORQKSLCVTAALVAGALAGVTIFSPVAILAGIVCGCLTALIQAFWQG 240
Db 181 LPALFMSFLLASFORQKSLCVTAALVAGALAGVTIFSPVAILAGIVCGCLTALIQAFWQG 240

Qy 241 APDEL 245
Db 241 APDEL 245

RESULT 3
Q8X908 ECO57 PRELIMINARY; PRT; 245 AA.
AC Q8X908; Q7ABD6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein EC83544.
GN OrderedLocusNames=EC83544, z3983;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005174; AAG57791.1; -; Genomic_DNA.
DR EMBL; BA000007; BAB36967.1; -; Genomic_DNA.
DR PIR; C85916; C85916.
DR PIR; H91071; H91071.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR011606; AzIC-like.
DR InterPro; IPR012294; TFIID_C/glycos_N.
DR Pfam; PF03591; AzIC; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 245 AA; 26093 MW; 20AAF2D91E644DB3 CRC64;

Query Match 99.9%; Score 1247; DB 2; Length 245;
Best Local Similarity 99.6%; Pred. No. 9.2e-93;
Matches 244; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESPTQPAPGSAATFMEGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIYAG 60

```

Db 1 MESPTQPAPGATFMGCKDSLPIVISYIPVAFAGLNATRLGSPLESVFSCIIYAG 60
 QY 61 ASQFVITAMLAAGSSLIWIAALTVMAMDVHRHLYGSPLSRIIQRLOKSKTALWAFGLTDE 120
 Db 61 ASQFVITAMLAAGSSLIWIAALTVMAMDVHRHLYGSPLSRIIQRLOKSKTALWAFGLTDE 120
 QY 121 VPAATAKLVRNRRWSENWMIIGIAFSSWSWVFGTVIGAFSGGSLGQYPAVEAALGFM 180
 Db 121 VPAATAKLVRNRRWSENWMIIGIAFSSWSWVFGTVIGAFSGGSLGQYPAVEAALGFM 180
 QY 181 LPALFMSFLLASFORKQSLCVTAALVGLAGVTLFSPVAILAGIVCCCLTALIOAFWQG 240
 Db 181 LPALFMSFLLASFORKQSLCVTAALVGLAGVTLFSPVAILAGIVCCCLTALIOAFWQG 240
 QY 241 APDEL 245
 Db 241 APDEL 245

RESULT 4

Q83JZ6_SHIFL
 ID Q83JZ6 SHIFL PRELIMINARY; PRT; 245 AA.
 AC Q83JZ6; Q7C0B7;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=S2896, SP2709;
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=623;
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=301 / Serotype 2a;
 RA MEDLINE=2227406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
 RT through comparison with genomes of *Escherichia coli* K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RX Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of *Shigella*
 RT flexneri serotype 2a strain 2457T.";
 RL Infect. Immun. 71:2775-2786(2003).
 DR EMBL; AE005674; AAN44202.1; -; Genomic DNA.
 DR EMBL; AE0016987; AAP18029.1; -; Genomic_DNA.
 DR GO; GO:0003677; P:DNA binding; IEA.
 DR InterPro; IPR011606; AzIC like.
 DR InterPro; IPR012294; TFIID C/glycos_N.
 DR Pfam; PF03591; AzIC; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 245 AA; 26079 MW; 3BAC440092FF6786 CRC64;

Query Match 99.6%; Score 1243; DB 2; Length 245;
 Best Local Similarity 99.2%; Pred. No. 1.9e-92;
 Matches 243; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESPTQPAPGATFMGCKDSLPIVISYIPVAFAGLNATRLGSPLESVFSCIIYAG 60
 Db 1 MESPTQPAPGATFMGCKDSLPIVISYIPVAFAGLNATRLGSPLESVFSCIIYAG 60

QY 61 ASQFVITAMLAAGSSLIWIAALTVMAMDVHRHLYGSPLSRIIQRLOKSKTALWAFGLTDE 120
 Db 61 ASQFVITAMLAAGSSLIWIAALTVMAMDVHRHLYGSPLSRIIQRLOKSKTALWAFGLTDE 120
 QY 121 VPAATAKLVRNRRWSENWMIIGIAFSSWSWVFGTVIGAFSGGSLGQYPAVEAALGFM 180
 Db 121 VPAATAKLVRNRRWSENWMIIGIAFSSWSWVFGTVIGAFSGGSLGQYPAVEAALGFM 180
 QY 181 LPALFMSFLLASFORKQSLCVTAALVGLAGVTLFSPVAILAGIVCCCLTALIOAFWQG 240
 Db 181 LPALFMSFLLASFORKQSLCVTAALVGLAGVTLFSPVAILAGIVCCCLTALIOAFWQG 240
 QY 241 APDEL 245
 Db 241 APDEL 245

RESULT 5

Q8ZBX0_YERPE
 ID Q8ZBX0_YERPE PRELIMINARY; PRT; 257 AA.
 AC Q8ZBX0; Q74W29; Q7CK78;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
 DE Putative amino acid transporter (Hypothetical protein y0925).
 GN Name=azlC; OrderedLocusNames=YPO668, YPO3264, Y0925;
 OS *Yersinia pestis*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Yersinia*.
 OX NCBI_TaxID=632;
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
 RA Leather S., Moulle S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
 RT "Genome sequence of *Yersinia pestis*, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RX DOI=10.1128/JB.184.16.4601-4611.2002;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of *Yersinia pestis* KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=91001;
 RX PubMed=15368893;
 RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
 RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
 RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
 RA Yang H., Wang J., Huang P., Yang R.;
 RT "Complete genome sequence of *Yersinia pestis* strain 91001, an isolate
 RT avirulent to humans.";
 RL DNA Res. 11:179-197(2004).
 DR EMBL; AJ414156; CAC92498.1; -; Genomic DNA.
 DR EMBL; AE013695; AAM84507.1; -; Genomic_DNA.
 DR EMBL; AE017129; AAS60335.1; -; Genomic_DNA.
 DR PIR; AF0396; AF0396.
 DR InterPro; IPR011606; AzIC like.
 DR Pfam; PF03591; AzIC; 1.
 KW Complete proteome; Hypothetical protein.

```
SQ SEQUENCE 257 AA; 27301 MW; 7C3660B0C8AD4FFPB CRC64;

Query Match 69.8%; Score 871.5; DB 2; Length 257;
Best Local Similarity 70.9%; Pred. No. 2.1e-62;
Matches 168; Conservative 29; Mismatches 35; Indels 5; Gaps 1;

QY 4 PTPQAPGS-----ATFMEGCKDSLPIVISYIPVAFAGLNATRLGSPLESVFFSCIIY 58
DB 2 ENPTAPSTTSAAATFVEGTDLSLPVIGVLPVAFAGLSAVKLGFTPLSFISFSCIIY 61
QY 59 AGAQFVITAMLAAGSSLIWIAALVTWMDVRHLYGSPLSRSRIIQRLOKSKTALWAFGLT 118
DB 62 AGAQFVITALLSAGMSLWVSALTVMMDVRHLYGSPALKHRIKLSGKKTALWAFGLT 121
QY 119 DEVFAAATAKLVRNRRSENWMIQIAFSSSSWVFGTVICAFSGSLGQYPAVEAALG 178
DB 122 DEVFAAATTKLMKQRRSENWMLGIAVTSLWSVLGTAIGAMFGNGPLENYPALAEASLS 181
QY 179 FMLPALFMSFLASFKRQYSITVIASLSGALLGVLLFSIPVAILAGIAGGCLATLIQ 235
DB 182 FMLPALFMSFLASFKRQYSITVIASLSGALLGVLLFSIPVAILAGIAGGCLATLIQ 238

RESULT 6
ID Q6E34_YERPS PRELIMINARY; PRT; 257 AA.
AC Q6E34;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DE Putative LIV-E family branched chain amino acid exporter, large subunit.
GN Name=YgagZ; OrderedLocusNames=YPTB0859;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype 1;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O., Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L., Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C., Simonet M., Chenail-Francoise V., Souza B., Dacheux D., Elliott J.M., Dextise A., Hauser L.J., Garcia E.;
RA "Insights into the evolution of Yersinia pestis through whole-genome comparison with Yersinia pseudotuberculosis.";
RT Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
RL EMBL; BX936398; CAH2009.1.-.; Genomic_DNA.
DR InterPro; IPR011606; AzlC; 1.
DR Pfam; PF03591; AzlC; 1.
KW Complete proteome.
SQ SEQUENCE 257 AA; 27297 MW; 88078AFD3104E45A CRC64;

Query Match 69.8%; Score 870.5; DB 2; Length 257;
Best Local Similarity 70.9%; Pred. No. 2.6e-62;
Matches 168; Conservative 29; Mismatches 35; Indels 5; Gaps 1;

QY 4 PTPQAPGS-----ATFMEGCKDSLPIVISYIPVAFAGLNATRLGSPLESVFFSCIIY 58
DB 2 ENPTAPSTTSAAATFVEGTDLSLPVIGVLPVAFAGLSAVKLGFTPLSFISFSCIIY 61
QY 59 AGAQFVITAMLAAGSSLIWIAALVTWMDVRHLYGSPLSRSRIIQRLOKSKTALWAFGLT 118
DB 62 AGAQFVITALLSAGMSLWVSALTVMMDVRHLYGSPALKHRIKLSGKKTALWAFGLT 121
QY 119 DEVFAAATAKLVRNRRSENWMIQIAFSSSSWVFGTVICAFSGSLGQYPAVEAALG 178
DB 122 DEVFAAATTKLMKQRRSENWMLGIAVTSLWSVLGTAIGAMFGNGPLENYPALAEASLS 181
QY 179 FMLPALFMSFLASFKRQYSITVIASLSGALLGVLLFSIPVAILAGIAGGCLATLIQ 235
DB 182 FMLPALFMSFLASFKRQYSITVIASLSGALLGVLLFSIPVAILAGIAGGCLATLIQ 238
```

```
Db 182 FMLPALFMSFLASFKRQYSITVIASLSGALLGVLLFSIPVAILAGIAGGCLATLIQ 238

RESULT 7
ID Q6D1D9_ERWCT PRELIMINARY; PRT; 243 AA.
AC Q6D1D9;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
DE Putative amino acid transporter.
DE OrderedLocusNames=ECA3509;
GN Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahnia M., Pritchard L., Holden M.T.G., Hyman L.J., Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K., Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J., Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H., Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S., Salmund G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RA "Genome sequence of the enterobacterial phytopathogen Erwinia carotovora subsp. atroseptica and characterization of virulence factors.";
RT Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
RL EMBL; BX950851; CAG76407.1.-.; Genomic_DNA.
DR InterPro; IPR011606; AzlC_like.
DR Pfam; PF03591; AzlC; 1.
KW Complete proteome.
SQ SEQUENCE 243 AA; 25805 MW; B00PB8F1B40A2EF1 CRC64;

Query Match 67.7%; Score 845.5; DB 2; Length 243;
Best Local Similarity 70.3%; Pred. No. 2.6e-60;
Matches 163; Conservative 28; Mismatches 40; Indels 1; Gaps 1;

QY 5 TPQAPGSATFMEGCKDSLPIVISYIPVAFAGLNATRLGSPLESVFFSCIIYAGASQF 64
DB 6 TPAPTK-SASFREGVDSLPVIGVMPVAFAGNNAVKLGFTPLFGLFSCIIYAGASQF 64
QY 65 VITAMLAGSSLIWIAALVTWMDVRHLYGSPLSRSRIIQRLOKSKTALWAFGLTDEVPA 124
DB 65 VITALLSAGMSIWAALTVMMDVRHLYGSPALRRHIVQQLPRTKRTALWAFGLTDEVPA 124
QY 125 ATAKLVRRNRRSENWMIQIAFSSSSWVFGTVICAFSGSLGQYPAVEAALGFMLPAL 184
DB 125 AATRLAKDNRNSENWMIQVSLLAWSVLGTVIGAVFGNGPLDNYPAVEAALSFMLPAL 184
QY 185 FMSFLASFKRQYSITVIASLSGALLGVLLFSIPVAILAGIAGGCLATLIQ 236
DB 185 FLSFLASFKRQYSITVIASLSGALLGVLLFSIPVAILAGIAGGCLASLVNS 236

RESULT 8
ID Q7N779_PHOLL PRELIMINARY; PRT; 252 AA.
AC Q7N779;
DT 01-MAR-2004 (TremBLrel. 26, Created)
DT 01-MAR-2004 (TremBLrel. 26, Last sequence update)
DE Similar to unknown protein Ygag of Escherichia coli.
DE OrderedLocusNames=plu1279;
GN Photobacterium luminescens (subsp. laumondii).
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photobacterium.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TT01;
```


RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=9-941 / Biovar 1;
 RX PubMed=15805518; DOI=10.1128/JB.187.8.2715-2726.2005;
 RA Halling S.M., Peterson-Burch B.D., Bricker B.J., Zuercher R.L.,
 RA Qing Z., Li L.-L., Kapur V., Alt D.P., Olsen S.C.,
 RT "Completion of the genome sequence of Brucella abortus and comparison
 RT to the highly similar genomes of Brucella melitensis and Brucella
 RT suis";
 RL J. Bacteriol. 187:2715-2726 (2005).
 DR EMBL; AF118548; AAL32283.1; -; Genomic DNA.
 DR EMBL; AB017223; AAX75147.1; -; Genomic_DNA.
 DR InterPro; IPR011606; AzlC_like.
 DR Pfam; PF03591; AzlC; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 224 AA; 23597 MW; 1576239416786FDC CRC64;
 Query Match 24.7%; Score 308.5; DB 2; Length 224;
 Best Local Similarity 35.1%; Pred. No. 5.9e-17;
 Matches 78; Conservative 45; Mismatches 86; Indels 13; Gaps 6;
 Qy 15 FMEGCKDLPVIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAGASQ-EVITAMLAAG 73
 Db 8 PFHGARSAPILIGLVPVAFAGTAAGSQGLALDSTALSMFSGANQAFFLSAVASGL 67
 Qy 74 SSLMIAALTVMAMDVHVLGYSRSLRIQRLQSKTALWAFGLTDEVFAA---ATAKLV 130
 Db 68 PTIAIVAICAVA-SLRHILYGFVLRRLAGL--ASRLAFAGLGTDEVFAATLVNATEK-- 122
 Qy 131 RNNRRWSENWMIATFSSWSWVGTGTFVAFSGSLGQYPAVEAALGFMPLPALFMSFLL 190
 Db 123 ---SKPDGWIIFGLAFWISWAATFFGAWMNILOAQFLQSLDALHFAFPALFLGLVW 179
 Qy 191 ASFORKQSLCVTAALVAGALAGVTLPFIPVAILAGIVCGCLTA 232
 Db 180 VTSARNVPMVAARVAVMFCL-LNLPALAIPOGASAAALLA 220
 Qy 15 FMEGCKDLPVIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAGASQ-EVITAMLAAG 73
 Db 8 PFHGARSAPILIGLVPVAFAGTAAGSQGLALDSTALSMFSGANQAFFLSAVASGL 67
 Qy 74 SSLMIAALTVMAMDVHVLGYSRSLRIQRLQSKTALWAFGLTDEVFAA---ATAKLV 130
 Db 68 PTIAIVAICAVA-SLRHILYGFVLRRLAGL--ASRLAFAGLGTDEVFAATLVNATEK-- 122
 Qy 131 RNNRRWSENWMIATFSSWSWVGTGTFVAFSGSLGQYPAVEAALGFMPLPALFMSFLL 190
 Db 123 ---SKPDGWIIFGLAFWISWAATFFGAWMNILOAQFLQSLDALHFAFPALFLGLVW 179
 Qy 191 ASFORKQSLCVTAALVAGALAGVTLPFIPVAILAGIVCGCLTA 232
 Db 180 VTSARNVPMVAARVAVMFCL-LNLPALAIPOGASAAALLA 220
 RESULT 14
 Q8FYL5 BRUSU
 ID Q8FYL5 BRUSU PRELIMINARY; PRT; 224 AA.
 AC Q8FYL5; BRUSU
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 23, Last annotation update)
 DE AzlC family protein.
 GN OrderedLocusNames=BR1853;
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OC NCBI_TaxID=29461;
 RN [1]
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Nayana L.A., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
 RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
 RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,
 RA Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
 DR EMBL; AE014291; AAN30748.1; -; Genomic_DNA.
 DR TIGR; BR1853; -;
 DR InterPro; IPR011606; AzlC_like.
 DR Pfam; PF03591; AzlC; 1.
 KW Complete proteome.
 SQ SEQUENCE 224 AA; 23597 MW; 1576239416786FDC CRC64;
 Query Match 24.7%; Score 308.5; DB 2; Length 224;
 Best Local Similarity 35.1%; Pred. No. 5.9e-17;

Matches 78; Conservative 45; Mismatches 86; Indels 13; Gaps 6;
 Qy 15 FMEGCKDLPVIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAGASQ-EVITAMLAAG 73
 Db 8 PFHGARSAPILIGLVPVAFAGTAAGSQGLALDSTALSMFSGANQAFFLSAVASGL 67
 Qy 74 SSLMIAALTVMAMDVHVLGYSRSLRIQRLQSKTALWAFGLTDEVFAA---ATAKLV 130
 Db 68 PTIAIVAICAVA-SLRHILYGFVLRRLAGL--ASRLAFAGLGTDEVFAATLVNATEK-- 122
 Qy 131 RNNRRWSENWMIATFSSWSWVGTGTFVAFSGSLGQYPAVEAALGFMPLPALFMSFLL 190
 Db 123 ---SKPDGWIIFGLAFWISWAATFFGAWMNILOAQFLQSLDALHFAFPALFLGLVW 179
 Qy 191 ASFORKQSLCVTAALVAGALAGVTLPFIPVAILAGIVCGCLTA 232
 Db 180 VTSARNVPMVAARVAVMFCL-LNLPALAIPOGASAAALLA 220
 RESULT 15
 Q6LZP2 METMP
 ID Q6LZP2 METMP PRELIMINARY; PRT; 238 AA.
 AC Q6LZP2; METMP
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE AzlC related protein.
 GN OrderedLocusNames=MWP0582;
 OS Methanococcus maripaludis.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanococcaceae; Methanococcus.
 OC NCBI_TaxID=39152;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=S2 / LL;
 RX PubMed=15466049; DOI=10.1128/JB.186.20.6956-6969.2004;
 RA Hendrickson E.L., Kaul R., Zhou Y., Boyee D., Chapman P., Chung J.,
 RA Conway de Macario E., Dodsworth J.A., Gillett W., Graham D.E.,
 RA Hackett M., Haydock A.K., Kang A., Land M.L., Levy R., Lie T.J.,
 RA Major T.A., Moore B.C., Porat I., Palmeiri A., Rouse G.,
 RA Saenphimmachak C., Soell D., Van Dien S., Wang T., Whitman W.B.,
 RA Xia Q., Zhang Y., Larimer F.W., Olson M.V., Leigh J.A.;
 RT "Complete genome sequence of the genetically tractable
 RT hydrogeotrophic methanogen Methanococcus maripaludis";
 RL J. Bacteriol. 186:6956-6969 (2004).
 DR EMBL; BX957220; CAF30138.1; -; Genomic_DNA.
 DR InterPro; IPR011606; AzlC_like.
 DR Pfam; PF03591; AzlC; 1.
 KW Complete proteome.
 SQ SEQUENCE 238 AA; 26219 MW; 70E25B7BBB74D035 CRC64;
 Query Match 24.2%; Score 301.5; DB 2; Length 238;
 Best Local Similarity 32.6%; Pred. No. 2.3e-16;
 Matches 78; Conservative 46; Mismatches 96; Indels 19; Gaps 6;
 Qy 15 FMEGCKDLPVIVISYIPVAFAGLNATRLGFSPLSFVFFSCIIYAGASQ-EVITAMLAAGS 74
 Db 10 YFSGIKDAIPISIGLPGVAFAGVLAWSMGIPVEISILMSLIIIFAGASQFVGNVLIAGT 69
 Qy 75 SLMIAALTVMAMDVHVLGYSRSLRIQRLQSKTALWAFGLTDEVFAAATAKLVNRNR 134
 Db 70 SSPEIVLTITLNLRLHFLMSSSLRSQRIDYTKSKKFLSLISFGVTDETFAVASLK---EEL 126
 Qy 135 RWSNWMIGTAFSSWSWVGTGTFVAFSGSLGQYPAVEAALGFMPLPALFMSFLASFP 193
 Db 127 KLSPEFLGLNFTAFPAWNPFTVLGIF---LAESIPKEIQSSMGISLYMFIQLLIPAV 182
 Qy 194 QRKQSL----CVTAALVAGALAGVTLPFIPVAILAGIVCGCLTALIQAFWOGAPDE 244
 Db 183 RRSSTKVLKFLVAVPFISSALTWPVPKFIETGTMIIITIVASFIGA---KFHGDGDE 238
 Search completed: February 15, 2006, 12:23:33

Job time : 207.02 secs

```

; APPLICANT: Gary L. Brston et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6525
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-6525

```

Query Match 47.4%; Score 592; DB 2; Length 266;
Best Local Similarity 48.1%; Pred. No. 2e-57;
Matches 111; Conservative 53; Mismatches 67; Indels 0; Gaps 0;

Qy	7	QAPG	SATF	WEGC	KDSI	PIV	ISY	Y	PVAF	AGLN	ATRI	GFSP	LESV	PSC	II	YAGASQ	PVI	66																						
Db	26	QMAA	QTAT	FQGA	KDSO	AAI	VL	TV	PSFA	FGVS	AQSF	QF	TPEW	EAFF	IL	SCSY	YAGASQ	FLV	85																					
Qy	67	TAMLA	AGSS	LI	WIAA	LT	WM	DV	RV	LV	XP	SL	RS	II	OR	LQ	SK	TAL	WAFGL	TV	DE	VFAA	AT	126																
Db	86	VALL	ASGS	SL	WT	ALT	VAL	D	IR	VL	YGP	AL	NY	LP	T	K	N	L	K	T	YV	W	ANG	L	TV	DE	VFA	SGM	145											
Qy	127	AKL	VR	NR	NR	RS	EN	WM	I	GA	T	SS	WS	S	VF	T	VG	I	AG	FS	G	L	L	Q	Q	Y	P	VA	EE	AL	GF	ML	PA	LF	186					
Db	146	IQL	SQ	RQ	QW	SE	SN	GL	S	LP	SW	MS	AG	SL	LG	L	F	AD	VA	HL	P	K	FL	QA	AL	D	FL	L	PA	L	FL	PA	L	FL	205					
Qy	187	SFL	LA	S	FOR	Q	S	L	C	V	T	AA	L	V	GA	L	A	G	V	T	L	F	S	P	V	A	I	A	L	G	V	C	C	T	A	L	I	QA	F	237
Db	206	SFL	LA	S	FOR	Q	S	L	C	V	T	AA	L	V	GA	L	A	G	V	T	L	F	S	P	V	A	I	A	L	G	V	C	C	T	A	L	I	QA	F	256

RESULT 3
US-09-489-039A-8478
; Sequence 8478, Application US/09489039A
; Patent NO. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8478
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8478

	Query Match	19.8%;	Score 246.5;	DB 2;	Length 247;
	Best Local Similarity	28.6%;	Pred. No. 4.5e-19;		
	Matches 67;	Conservative 46;	Mismatches 106;	Indels 15;	Gaps 6;
Qy	18	GCKDSLPIVSIYIVAFAGFLNATRLGFSPLSEVFFSCIIYAGASQFVITAMLAGGSLW	77		
Db	10	GASAIMPLCIGDPFSTFIVGALSVSAGSVWQSTAWSAIVTAGSQMLALNMLKTGATLG	69		
Qy	78	IAALTVMAMDVHVLYGSLRSRIIQRLQKSKTALWAFGLTDEVPAAATAKLVNRNRWS	137		
Db	70	VIIFTTLIIINLRHVLVSASI-SGTVREASPFKCKFMSYALTDEVY-ATTVKEMGCKKEK	127		
Qy	138	ENWMIGIAFSWSWVFCTVIGATSGGSLQGYPAVEA-ALGEMLPALPMSFLIASFORK	196		
Db	128	YLFYGSAMITFWALWLDADFALGVGA----SFPHIEKYGDFPWAAFAIVVP--QIK	181		
Qy	197	QSLCVTAALVGALAGVTILFSP-----VAIAGIVCGCLTALTQIAFWQGPAPDE	244		
Db	182	SOACTVAAVVAASGVLLVLPYSILGIWASVLGVLAGLCVYDLAERKOMAKTE	235		

```

RESULT 4
US-09-710-279-268
; Sequence 268, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 268
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-268

```

	Query Match	18.4%;	Score 229.5;	DB 2;	Length 230;
	Best local Similarity	28.6%;	Pred. No. 3.2e-17;		
	Matches 65;	Conservative 47;	Mismatches 102;	Indels 13;	Gaps 6;
Qy	14	TFMEGCKDSLPIVTSYIPVAFAGFNATRLGFSPLSFVFFSCIIYAGASQRFVITAMLAAG	73		
Db	5	TFKGVKECTPTLLGYAGVGLSGFIVASQNFVLEILLCLIIYAGAQFIICTLVIAG	64		
Qy	74	SSLWIAALTVMMDVRHVLGYPSLRRIIQRLQKSKTALWAFGLTDEVFAAATAKLVRNN	133		
Db	65	TPISAIVLITLIVNSRMFLLSMTLAPKNYQYGFNNRVGLGTL-LTDETFGVAITPYVK-G	122		
Qy	134	RWSENWMIQAFSSWSWSPGTVIGAFSGSGLLQGYPAVEAAG--FMLPALFMSPLLA	191		
Db	123	EKINDRDLHGNIITAYLFWTVSCVIGAFGE-----YISNPDALGLDFAITMFIICIS	177		
Qy	192	SFQ--RQKSLCVTAALYGALAGVTLF--SIPVAILAGIVCGCLTALI	234		
Db	178	OFEGIGKSRRIYITVLVGVVMMLLSLIPSYLAILIAIAVAALL	224		

RESULT 5
US-09-134-001C-5154
; Sequence 5154, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: LYNN DOUCETTE-STAMM ET AL
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING
; TO THE IDENTIFICATION OF STAPHYLOCOCCUS EPIDERMIDIS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5154
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5154

Query Match	18.1%	Score 225.5;	DB 2;	Length 242;
Best Local Similarity	28.6%;	Pred. No. 9.5e-17;		
Matches	66;	Conservative 39;	Mismatches 89;	Indels 37; Gaps 7
Qv	14	TFMEGCKDSLPIVISTYIPVAFAGLNATRLGSPLESVFPSCLTYAGASQFVITAMLAAG	73	

[illegible]

```

; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4486
; LENGTH: 263
; TYPE: PRF
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-4486

Query Match          16.0%; Score 199.5; DB 2; Length 263;
Best Local Similarity 29.0%; Pred. No. 8.3e-14;
Matches 65; Conservative 41; Mismatches 101; Indels 17; Gaps 6;

Qy 15 FMECKDQLPIVIGSYVPAPAFGLNATRLGSPLESVPFFSCIIYAGASQFVITAMLAAGS 74
Db 45 FLRGAILDLPUSIVIPWAILAGSNVHAGLSFYKALAMSGIVPAGAAQLVLSNWEGA 104

Qy 75 SLWTAALTVMAMDVRHVLVYGPSLSRIIQRLOKSKTALWAFGLTDEVFVFAATAKLVRNNR 134
Db 105 SLLTIYVTFIPLTAQHFIYALTLEND-LSILPLSKRLTGLLLTDELFAVSVSP-----NE 158

Qy 135 RWSENNWIGIAFSSWSNVFQTVIGAFSGGL--LQGYPAVEAALGFMLPALFMSFLLAS 192
Db 159 KCHPQYLFAGLCFPYFWVPSLVGILLATLAPLNLY----HLDFSIITAIFFVAMIVPM 213

Qy 193 FQRQSKLCVTHAALVGAAGVTL--FSIPVAILAGIVCGCLTALI 234
Db 214 CKGKXP--VMAGILMTCVSGFVFLKPFHIEGAILISGLLGMFITAVI 255

```

RESULT 9
 US-09-471-803A-3
 ; Sequence 3, Application US/09471803A
 ; Patent NO. 6613545
 ; GENERAL INFORMATION:
 ; APPLICANT: KENNERNECHT, NICOLE
 ; APPLICANT: SAHM, HERMANN
 ; APPLICANT: EGELING, LOTHAR
 ; APPLICANT: PFEFFERLE, WALTER
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
 ; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
 ; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
 ; FILE REFERENCE: 21123/265496/MAS
 ; CURRENT APPLICATION NUMBER: US/09/471,803A
 ; CURRENT FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: DE 199 51 708.8
 ; PRIOR FILING DATE: 1999-10-27
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 251
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; OTHER INFORMATION: ATCC14752
 US-09-471-803A-3

Query Match	15.8%;	Score 197;	DB 2;	Length 251;
Best Local Similarity	27.8%;	Pred. No. 1.5e-13;		
Matches 65;	Conservative 47;	Mismatches 92;	Indels 30;	Gaps 9;
Qy	17	EGKDSLPVISYTPVAPAFGLNATRLGFSPLESVPFSCIIYAGASQFVITAMLAAGSSL	76	
Db	34	QGLKTSLAAGLGMPYPIGIAFLGLVIQGYEWAAAPLFSGLI FAGSTEMLVIALVVGAAPL	93	
Qy	77	WIAALTVMAMDVRHVLGSPLSRRIIOKLQSKTALW--AFGLTDEVFAAATAKLVRNRR	134	
Db	94	GAIALTLLVNFRRHVFYAFSPPLHVV---KNPIARFYFVALIDEAVAVTAAR----	145	PA
Qy	135	RWSENWIGIAFSWSWVFCTVIGAPSGGLLOQYPAVEAALGFMLPAFWSFLLASFQ	194	
Db	146	GWSAWRLISMQIAFHSYVWF---GGLTGVIAIELIPFEIKGLEFALCSLEFVTLTDSR	201	
Qy	195	RKQSLCVTAALVAGALGVTLSFIPVAILAG-----IVCGCLTALIQAFWQG	240	

Db 202 TKKQI----PSLL--LAGLS-FTIALVVPQALPAALLIFGLLT--IRYFFLG 247

RESULT 10
US-10-608-504-3
; Sequence 3, Application US/10608504
; Patent No. 6841360
; GENERAL INFORMATION:
; APPLICANT: KENNERNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGLING, LOTHAR
; APPLICANT: PEPPERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 7601/80525
; CURRENT APPLICATION NUMBER: US/10/608,504
; PRIOR FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: US 09/471,803
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; OTHER INFORMATION: ATCC14752
US-10-608-504-3

Query Match	15.8%;	Score 197;	DB 2;	Length 251;
Best Local Similarity	27.8%;	Pred. No. 1.5e-13;		
Matches	65;	Conservative 47;	Mismatches 92;	Indels 30; Gaps 9;
Qy	17	EGCKDSLPIVTSYTPVAFAGFNATRLGFSPLSEVFVSCIIYAGASOFVITAMLAAGSSL	76	
Db	34	QGLKTSLAAGLWYPIGATGELLVIQGYEWAAPLFSGLIFAGSTEMVLVALVGAAPL	93	
Qy	77	WIAALTVMAMDVRHVLVGPSLRSRIIORLOKSKTALM--AFGLTDEVFAAATAKLVRNNR	134	
Db	94	GAIALTTLLNVNFRHVFAFSPPLHW---KNPIARFVSVPALIDEAYAVTAAR----	145	PA
Qy	135	RWSENWMIAGTAFSSWSWVGCTVTAFGSGCLLOGYPAVEAALGFMPLPALFMSFLASFO	194	
Db	146	GWSAURLISMQIAFHSYWFV---GGTGVNAIAELIPEIKGLSEFALCSLFEVITLDSCR	201	
Qy	195	RKQSLCVTAALVAGLAGVTLFSIPVALAG-----IVCGCITALIQAQWOG	240	
Db	202	TKKQI-----PSLIL---LAGLS-PTIALVVTIPGOALFAALLIFLGILT---IRYFELG	247	

```

RESULT 11
US-09-605-703B-1840
; Sequence 1840, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605.703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934

```

```
; SEQ ID NO 1840
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-1840

Query Match      15.8%; Score 197; DB 2; Length 251;
Best Local Similarity 27.8%; Pred. No. 1.5e-13;
Matches 65; Conservative 47; Mismatches 92; Indels 30; Gaps 9;

Qy 17 EGCKDSLPIVISYIPVAFAGLNATRLGSPLESVFSCIIYAGASQFVTITAMLAAGSSL 76
Db 34 QGLKTSLAAGLGMTPIGIAFGLLIQGYEWAAPLFSLGILFAGSTEMLVIALVVGAAPL 93
Qy 77 WIAALTVMAMDVHRVLYGPSLRSRRIORLQSKTALW--AFGLTDEVFAAATAKLVRNNR 134
Db 94 GAIALTLLVNVFRHVFVAFSPPLHV---KNPIARFVSVPALIDEAYAVTAAR----PA 145
Qy 135 RWSNNWIGIAFSSWSNVFVTGAFSGGLQGYPAVEAALGFMLPALFMSFLASFO 194
Db 146 GWSAWRLISMQIAPHSYWVF---GGLTGVAIAELIPFEIKGLEFALCSLFTVLTLSR 201
Qy 195 RKQSLCVTAALVAGLAGVTLFSPVAILAG-----IVCGCLTALIQAFWQ 240
Db 202 TKKQI---PSLL--LAGLS-FTIALVPIQALFAALLIFLGLT--IRYFFLG 247

RESULT 12
US-09-134-000C-6789
; Sequence 6789, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6789
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6789

Query Match      15.6%; Score 194.5; DB 2; Length 206;
Best Local Similarity 31.1%; Pred. No. 2.1e-13;
Matches 65; Conservative 29; Mismatches 92; Indels 23; Gaps 7;

Qy 44 GSPLESVFSCIIYAGASQFVTITAMLAAGSSLWIAALTVMAMDVHRVLYGPSLRRI-I 102
Db 7 GSPSLIVSAMSFFIFAGSAQFVTVSMLTGGSPILSVLATFLVNARMILGMGTIAPYFKA 66
Qy 103 QRLQSKTALWAFQ--LTDEVFAAATAKLVRNNRWSNNWIGIAFSSWSNVFRTGIGA 160
Db 67 ESIGKN---LW-LGTLTDSFALGMNKLHNTKRLSFEWNAANLISYATWVFSTIIGA 122
Qy 161 FSGSGLLQGYPAVEAALG--FMLPALFMSFLASFORQSLCVTAALVAGLAGVTLFSP 218
Db 123 YLGR-----FIANPQALGLEFAVAVMFIQGLLYLQIISDRSMKIALQJLWLTIFGLMSIG 177
Qy 219 -----VAIAGIVCGCLTALIQAFW 238
Db 178 LIFIPSNLIVLVVTTLGCGIRVMIKHAF 206

RESULT 13
US-09-583-110-4834
; Sequence 4834, Application US/09583110
; Patent No. 6699703

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4834
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (188)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-583-110-4834

Query Match      15.0%; Score 187; DB 2; Length 224;
Best Local Similarity 25.6%; Pred. No. 1.6e-12;
Matches 60; Conservative 48; Mismatches 96; Indels 30; Gaps 10;

Qy 17 EGCKDSLPIVISYIPVAFAGLNATRLG---FSPLESVFSCIIYAGASQFVTITAMLAAG 73
Db 1 EGAQAAMPTALGYVSGIACG---TIGAPYVTPVEMGLMSLFYVYAGSAQFAMLLIIVVQ 56
Qy 74 SSIWIAALTVMAMDVHRVLYGPSLRRIORLQSKTALW-----AFGLTDEVFAAATAK 128
Db 57 APVAALATMTVFLINLR--LFLLSHASTYPR----HTSLWNMGMSILLIDETTVGLMGE 110
Qy 129 LVNRNRWSNNWIGIAFSSWSNVFRTGAFSGGLQGYPAVEAALGFMLPALFMSF 188
Db 111 LAHTD-KVNDPMWGHNNLSYVAVFWGVTVGTALG-GLLPN-PEI-FGLDFALVGMFIGI 166
Qy 189 LLASFQKQS-----LCVTAALVAGLAGVTLFSPVAILAGIVCGCLTALI 234
Db 167 FASQFQMMORRIPVRNLLIILXVAVSVYFLLLTWVSQSLAVLFAMLLGCSMGVV 220

RESULT 14
US-09-605-703B-2342
; Sequence 2342, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: EGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 2342
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-2342

Query Match      14.4%; Score 180; DB 2; Length 237;
Best Local Similarity 27.8%; Pred. No. 1.1e-11;
```

Matches 59; Conservative 39; Mismatches 88; Indels 26; Gaps 6;

Qy 18 GCKDSLPIVISYIPVAFAGLNATRLGFSPLSFSCIIYAGASQFVITAMLAAGSLW 77

Db 12 GIGETLTVGLGLIPLGLAFGLLWQTGFANWTFISFVIYAGSMEFLAIGMVTAGICPF 71

Qy 78 IAAALTVMAMDVHVLGPSL-RSRIIQRLQSKT--ALWAFGLTDEVFAAATAKLVRNR 134

Db 72 SAAVAGFVNPRHFYGLTFFPRHRI-----KSGAGRAYSTYALTDESAIVSA---RPPG 123

Qy 135 RWSNNMIGIAFSSWSSWVFTVIGAFSGGLLQCYPAVEAALGFMPLPALFMSFLLASFQ 194

Db 124 DISGTRVLTVQLCOALWVIGIIGALVG-----QVLPDDLKGMDFALTALFVVLAWAEAFK 179

Qy 195 RKQSLCTVAALVGLAGLVTLFSIPVAILAGIV 226

Db 180 NKKDY-----SLPLFAVVLALVSGFV 200

RESULT 15

US-09-107-433-3341

; Sequence 3341, Application US/09107433

; Patent No. 680744

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO

; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: <Unknown>

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: <Unknown>

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,433

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/ 085131

; FILING DATE: May 12, 1998

; APPLICATION NUMBER: 60/051553

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-011

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 3341:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 185 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Streptococcus pneumoniae

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...185

; SEQUENCE DESCRIPTION: SEQ ID NO: 3341:

US-09-107-433-3341

Query Match 13.7%; Score 171; DB 2; Length 185;

Best Local Similarity 28.0%; Pred. No. 7.4e-11;

Matches 53; Conservative 38; Mismatches 76; Indels 22; Gaps 9;

Qy 17 EGCKDSLPIVISYIPVAFAGLNATRLG---FSPLESVFFSCIIYAGASQFVITAMLAAG 73

Db 1 EGAQAAMPTALGYVSIIGLACG---TTIGAPYVTVPEVMGLMSLFYVYAGSAQFAMLALIVVQ 56

Qy 74 SSLWIAALTVMAMDVHVLGPSLRSRIIQRLQSKTALW-----AFGLTDEVFAAATAK 128

Db 57 APVAAIAIANTVFLINLR--LFLLSLHASTYFR---HTSLWNINIGMSSILTDITYGVLME 110

Qy 129 LVNRNREWSNNMIGIAFSSWSSWVFTVIGAFSGGLLQCYPAVEAALGFMPLPALFMSF 188

Db 111 LAHTD-KVNPMPMWHGNNLSYVAMFVGTVVGTTALG-GLLPN-PEI-FGLDFALVGMFTGI 166

Qy 189 LLASDFQRKQ 197

Db 167 FASQFQMMQ 175

Search completed: February 15, 2006, 12:25:38
Job time : 47.1096 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 12:24:42 ; Search time 157.598 Seconds
(without alignments)
649.551 Million cell updates/sec

Title: US-10-073-293A-4
Perfect score: 1248
Sequence: 1 MESPTQPAPGSAFMFGCK.....VCGCLTALIOAFMQGAPDEL 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgm2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgm2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1248	100.0	245	US-10-073-293A-4	Sequence 4, Appli
2	225.5	18.1	242	US-10-724-972A-6721	Sequence 6721, Ap
3	197	15.8	251	US-09-738-626-3789	Sequence 3789, Ap
4	197	15.8	251	US-10-608-504-3	Sequence 3, Appli
5	195.5	15.7	229	US-10-335-977-6207	Sequence 6207, Ap
6	189	15.1	226	US-10-501-282-2600	Sequence 2600, Ap
7	182.5	14.6	297	US-10-156-761-10733	Sequence 10733, A
8	180	14.4	237	US-09-738-626-6908	Sequence 6908, Ap
9	178	14.3	218	US-10-474-776-217	Sequence 217, App
10	178	14.3	218	US-10-472-928-96	Sequence 96, Appl
11	172	13.8	214	US-10-501-282-2598	Sequence 2598, Ap
12	171	13.7	185	US-10-617-320-3341	Sequence 3341, Ap
13	153.5	12.3	302	US-10-156-761-10035	Sequence 10035, A
14	105	8.4	324	US-09-738-626-4887	Sequence 4887, Ap
15	102.5	8.2	474	US-09-738-626-6653	Sequence 6653, Ap
16	102	8.0	301	US-10-282-122A-45235	Sequence 45235, A
17	100	8.0	357	US-10-282-122A-46768	Sequence 46768, A
18	99.5	8.0	307	US-10-282-122A-66244	Sequence 66244, A
19	98.5	7.9	467	US-09-991-212A-3	Sequence 3, Appli
20	98.5	7.9	467	US-09-965-522-3	Sequence 3, Appli
21	98.5	7.9	467	US-10-877-818-3	Sequence 3, Appli
22	98	7.9	432	US-10-369-493-13811	Sequence 13811, A
23	98	7.9	535	US-10-282-122A-43404	Sequence 43404, A
24	98	7.9	535	US-10-282-122A-60010	Sequence 60010, A
25	97.5	7.8	381	US-09-738-626-3549	Sequence 3549, Ap
26	97.5	7.8	441	US-09-950-071-2	Sequence 2, Appli
27	97.5	7.8	441	US-10-721-922A-116	Sequence 116, App

ALIGNMENTS

RESULT 1

US-10-073-293A-4
; Sequence 4, Application US/10073293A
; Publication No. US20050239175A1
; GENERAL INFORMATION:
; APPLICANT: TABOLINA, EKATERINA
; APPLICANT: RYBAK, KONSTANTIN
; APPLICANT: KHOURGES, EVGENI
; APPLICANT: VOROSHILOVA, ELVIRA
; APPLICANT: GUSYATINER, MIKHAIL
; TITLE OF INVENTION: METHOD FOR PRODUCING L-AMINO ACID USING BACTERIA BELONGING TO
; TITLE OF INVENTION: ESCHERICHIA
; FILE REFERENCE: 21594USO
; CURRENT APPLICATION NUMBER: US/10/073,293A
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: RU 2001103865
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: RU 2001104998
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: RU 2001104999
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: RU 2001117632
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: RU 2001117633
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-073-293A-4

Query Match 100.0%; Score 1248; DB 5; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.4e-117;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESPTQPAPGSAFMFGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIYAG 60
Db 1 MESPTQPAPGSAFMFGCKDSLPIVISYIPVAFAGLNATRLGFSPLSFVFFSCIYAG 60

Qy 61 ASQFVTMLAAGSSLWIAALTVMADVRHVLVGPSLRSRIIORLQSKTALWAFGLTDE 120
Db 61 ASQFVTMLAAGSSLWIAALTVMADVRHVLVGPSLRSRIIORLQSKTALWAFGLTDE 120

Qy 121 VFAAATAKLVRRNRRSENNMIGIAFSSSSWVFGTVIGAFSGLLQGYPAVEAALGFM 180
Db 121 VFAAATAKLVRRNRRSENNMIGIAFSSSSWVFGTVIGAFSGLLQGYPAVEAALGFM 180

Qy 181 LPALFMSFLIASPQRKQSLCVTAALVAGALAGVTLSIPVAILAGIVCGCLTALIQAFWQ 240

Sequence 44923, A
Sequence 371, App
Sequence 49508, A
Sequence 77832, A
Sequence 47464, A
Sequence 46646, A
Sequence 3, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 77660, A
Sequence 61304, A
Sequence 3877, Ap
Sequence 3212, Ap
Sequence 23813, A
Sequence 299, App
Sequence 1978, Ap
Sequence 4191, Ap
Sequence 14, Appli

```

|||||
Db 181 LPALFMSGLLASFORKOSLCVTAALVGLAGVTLFSIPVAILAGVCGCLTALIQAFWQG 240
|||
Qy 241 APDEL 245
|||
Db 241 APDEL 245

RESULT 2
US-10-724-972A-6721
; Sequence 6721, Application US/10724972A
; Publication No. US20040147734A1
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-16
; CURRENT APPLICATION NUMBER: US/10724.972A
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/450,969
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 6721
; LENGTH: 242
; TYPE: PRT
; ORGANISM: S.epidermidis
US-10-724-972A-6721

Query Match 18.1%; Score 225.5; DB 4; Length 242;
Best Local Similarity 28.6%; Pred. No. 3.2e-14;
Matches 66; Conservative 39; Mismatches 89; Indels 37; Gaps 7;

Qy 14 TFMGGKDSLPIVSYIPVAFPAFGLNATRLGFSPLSVFFSCIIYAGASQFVITAMLAAG 73
Db 17 TFKQGVKECIPITLGYAGVGLSPGIVAAQNFVSLEILLCLIIYAGAAQFIICLTVIAG 76
Qy 74 SSLWIAALTVMAMDVRRHVLVGPSSRSRIORLQKSKTALMAFGLTDEVFAAATAKLVRNN 133
Db 77 TPISAIVLTLIVNSRFLFSLNTLAPNYKQGFNVRVGLGTL-LDETFGVAITPYVK-G 134
Qy 134 RRWSENWMIAGFSWSWSWVGTVIGAFSGSGLLQGYPAVEAALG--FMLPALFMSFLLA 191
Db 135 EKINDRWLHGLNITAYLFTWTVSCVIGAIFGE-----YISNPDALGLDPAITAMRIFLCIS 189
Qy 192 SFQ--RKQSL-----CV-----TAALVGLAGVTL 214
Db 190 QFEGIKESRLRIYIVLIVCVIWMMLLSILPYSVAILIAAIVAAALLGVWM 240

RESULT 3
US-09-738-626-3789
; Sequence 3789, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125

```


Qy 77 WIAALTVMAMDVRLVYGPSLRSRIORLQSKTALW--AFGLTDEVFAAATAKLVRNR 134
Db 94 GAIALTILLVNRFRVFAFSPPLHV-----KNPIARFYSFALIDEAYAVTAAR----PA 145
Qy 135 RWSNWMGIAFSSWSWVFTVIGAFSGSLGLOGYPAVEAALGFMLPALFMSFLLASFQ 194
Db 146 GWSAWRLSMGIAFHSYWF-----GGLTGVAIAELIPPEIKGLEFALCSLFVTLTLDSCR 201
Qy 195 RKQSLCVTAALVAGALAGVTLFSIPVAILAG-----IVCGCLTALIQAFWOG 240
Db 202 TKQOI---PSLL--LAGLS-FTIALVVIPOALFAALLIFLGLT--IRYFFLG 247

RESULT 5
US-10-335-977-6207
; Sequence 6207, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335.977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...229
; SEQUENCE DESCRIPTION: SEQ ID NO: 6207:
US-10-335-977-6207

Query Match 15.7%; Score 195.5; DB 4; Length 229;
Best Local Similarity 27.0%; Pred. No. 3.2e-11;
Matches 62; Conservative 42; Mismatches 85; Indels 41; Gaps 9;
Qy 15 FMEGCKDSLPIVISYIPVAFAPGLNATRLGFSPLSFVFCIIYAGASQFVITAML 70
Db 5 FLKAFKADPHTTISIFIGYLLMGMTFGMLLAQQQDYKVALFMSLFIYAGAIQFVAILL 64
Qy 71 AAGSSLIAALTVMAMDVRLVYGPSLRSRIORLQSKTALW-----AFGLTDEVFAA- 124

Db 65 SAQASLMNVVIVSLVNNARQTCYALSM-----LDRFKNTKWLRYLAHALTDETFA 117
Qy 125 ---ATAKLVRNRWSNWMGIAFSSWSWVFTVIGAFSGSL---LQGYPAVEAALG 178
Db 118 NLVAPKGVNET-----DFMFSILLNHSWTWPSGLVSGLSVGHFSFDTQ3-----ME 165
Qy 179 FMLPALFMSFLLASFQKQS-----LCVTAALVG-ALAGVTLFSIPVAIL 222
Db 166 FVMTAIFVLFMQYKNTNHNKAWLGIAIAVAVCLALFGTEYFLIALVL 215

RESULT 6
US-10-501-282-2600
; Sequence 2600, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCCOCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2600
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Alloiooccus otitidis
US-10-501-282-2600

Query Match 15.1%; Score 189; DB 5; Length 226;
Best Local Similarity 23.9%; Pred. No. 1.4e-10;
Matches 55; Conservative 52; Mismatches 99; Indels 24; Gaps 6;
Qy 15 FMEGCKDSLPIVISYIPVAFAPGLNATRLGFSPLSFVFCIIYAGASQFVITAML 74
Db 2 FKKVLKPAFPVMVSYLVSLVCGIVSFQVGFPLQILLTSVLYSGSQFLLAGLYGAGA 61
Qy 75 SLMIAALTVMAMDVRLVYGPSLRSRIORLQSKTALWAF-----GLTDEVFAAATAKL 130
Db 62 SLVSIITLTAFLGLRFLVMSSSSRHV-----RQKTTWDFDFFSMTISDESFGVNTWFS 116
Qy 131 RNNRWSNWMGIAFSSWSWVFTVIGAFSGSLGLOGYPAVEAALGFMLPALFMSFLL 190
Db 117 QPD--WTADHALALNLLNYGIWLGSLGAL-----LVSVVDLDTSIISYGLTAMFCMT 170
Qy 191 ASFORKQSLCVTAALVAGALAGVTLFSIPVAIL---AGIVCGCLTALIQAF 237
Db 171 EQFVDY-----LYAGLSIVVFTTIALVILQNSLGIYVVGALLASLIGF 214

RESULT 7
US-10-156-761-10733
; Sequence 10733, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA

```

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10733
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10733

```

Query Match	14.6%;	Score	182.5;	DB	4;	Length	297;
Best Local Similarity	25.5%;	Pred.	No. 9e-10;				
Matches	65;	Conservative	47;	Mismatches	120;	Indels	23; Gaps
							8;
Qy	7	QPAPGSATFMEGCKDSLPIVISYIPVAFAGLNATRIGFSPLESVFSSFCIIYAGASQFVI	66				
Db	16	RPGPGS-TLRAGAMSLPFLASAIPMGLAGGILGPANGLSGWATLGLAMAVNSCTAQFIG	74				
Qy	67	TMLAAGSSLWIAALTWMADVHVLCPSLRSRIIQRLQSKSTALWAGLTIDVEFAANT	125				
Db	75	FALIAGHGSAITILLTTLIIGRLMIYSAILPHVREVPQRVLL-GFGILDIVFFVAI	133				
Qy	127	AKLVRRNRRWSEN-WMTIGIAPSSWSNVGTGVIGAFSGSLGLOGYPAVEAALGFMLPALF	185				
		: : : : : : : : : : : : : : : : : :					
Db	134	ERLKKGELTERKHLFPFGASGMVVTWNMLCTLVGMALGSAVPD---LADLGLDPMTAMF	190				
		: : : : : : : : : : : : : : : : : :					
Qy	186	MSFLIASFO--RKQSLCVTAALVGALA-----GVTLFSIPVAILAGIVCGCI-----	230				
Db	191	VAMLAGSLANWRVGAALVTAGLTVVVLGHGLPYNLGVVLATV-VGATGTVCVEYLKERTP	249				
Qy	231	-TALIQAFWQGAPDE	244				
		: : : : :					
Db	250	ENSATETSBESSPDE	264				

```

RESULT 8
US-09-738-626-6908
; Sequence 6908, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6908
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6908

```

```

Query Match      14.4%; Score 180; DB 3; Length 237;
Best Local Similarity 27.8%; Pred. No. 1.2e-09;
Matches 59; Conservative 39; Mismatches 88; Indels 26; Gaps 6;

QY 18 GKDSLPIVISYIPVAFAGLGNATRLGFSPLSFVSFSCIIYAGASOFVITAMLAAGSSLW 77
Db 12 GIGETLVGLGLPLGLAGLGMVQGFPAWWTPIFSFYIYAGSMFLAIGMTAGIGPF 71
QY 78 IAAALTVMAMDVRHVLXGPSL-RSRITQLQKSKT--ALMAFAGLTDEVFAAATAKLVRNNR 134
Db 72 SAAVAGFMWNFRHIFGLYFPFRHRI-----KSGAGRAYSTYALTDSSYALVSA--RPPG 123
QY 135 RWSENMMWIGAFSSWSWVFVTGIGAPSGGLLQGYPAVEAALGFMPLAFMSFLASFP 194
Db 124 DISGTRVLTVQILCOALWIPGIIIGALVG---QVLPDDLKGMDFALTALFVVLAWAEAFK 179
QY 195 RKOSLCVTAALVAGALAGVTLFSIPVAILAGIV 226
Db 180 NNNKDY-----SLPLFAVVLALVSGFV 200

RESULT 9
US-10-474-776-217
; Sequence 217, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCOD
; FILE REFERENCE: AM100649-PCT
; CURRENT APPLICATION NUMBER: US/10/474,776
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 217
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-217

```

```

Query Match      14.3%; Score 178; DB 4; Length 218;
Best Local Similarity 26.3%; Pred. No. 1.8e-09;
Matches 60; Conservative 47; Mismatches 91; Indels 30; Gaps 11;

Qy 23 LPIVISYIPVAFAPGLNATRLG---FSPLESVEFFSCIIYAGASQFVITAMLAAGSSLIWIA 79
Db 1 MPTALGVISGLACGI-----IGAPYVTPVEMGLMSLFPVYAGSAQFAMLAIVVQAPVAAI 56
Qy 80 ALVTAMDVRHVLYGSLRSRIIRLOQKSTALW-----AFGLTDEVFAAATAKTLVNNR 134
Db 57 AMTVFLINLR--LFLSLSHASTYFR---HTSLWYNIGMSSILTDTGYGLMGELAAHTD- 109
Qy 135 RWSENNMIAGFSSWSWVFGTVIGAFSSGSLQGYPAVEAALGFWMLPALFMSFLLASFO 194
Db 110 KVAMPMMHGNILNSYVAFVGTVGVGTALG-GLLPN-PEI-FGLDFALVGMFGIFAFASQFO 166
Qy 195 RKQ-----SLCVTAALVGA--LAGVTLFSIPVAIIAGIVCGCLTALI 234
Db 167 IMQRRIPVRNLLIILAVAVSFFLLLTVMSSQSLAVLFATLLGCGSMGV 214

RESULT 10
US-10-472-928-96
; Sequence 96, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926W0
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658-7

```

```

; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqMin99, version 1.03
; SEQ ID NO 96
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: conserved hypothetical protein
; OTHER INFORMATION: Cellular location: membrane
; OTHER INFORMATION: Similar to strain R6 sequence 15902189 (2.E-55)
US-10-472-928-96

Query Match      14.3%; Score 178; DB 5; Length 218;
Best Local Similarity 26.3%; Pred. No. 1.8e-09;
Matches 60; Conservative 47; Mismatches 91; Indels 30; Gaps 11;

Qy 23 LPVISYIPVAFAPAGLNATRLG---FSPLESVPFSCIYAGASQFVTMTMLAAGSSLIWA 79
Db 1 MPTALGYVSGIACGI---IGAPYVTPVEVGLMSLFYAGSAQFAMALITVQAPVAAI 56
Qy 80 ALTVMAMDMVRHVLGSPSLRSRIIQRLOKSKTALW----AFGLTDEVFAAATAKLVRNNR 134
Db 57 AMTVFLNLR--LFELSLHASTYFR----HTSLWNMGMSILTDYGVLMGSLAHTD- 109
Qy 135 RWSENMMIGIAPSSWSWVFTVIGAFSGGLQGYPAVEAALGFMLPALFMSFLLASFO 194
Db 110 KVPNMMHGNLNSYVAFVGVTVGTALG-GLLPN-PEI-FGLDFALVGMFEGIFASQFQ 166
Qy 195 RKQ-----SLCVTAALVGA--LAGVTLFSPVAILAGIVCGCTALAI 234
Db 167 IMORRIPVRNLLIILAVVAVSFLLITVMSQSLAVLPATLLGCSMGVV 214

RESULT 11
US-10-501-282-2598
; Sequence 2598, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID FARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCCOCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2598
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Alloiooccus otitidis
; US-10-501-282-2598

Query Match      13.8%; Score 172; DB 5; Length 214;
Best Local Similarity 23.7%; Pred. No. 6.9e-09;
Matches 52; Conservative 49; Mismatches 94; Indels 24; Gaps 6;

Qy 26 VISYIPVAFAPAGLNATRLGSPLESVPFSCIYAGASQFVTMTMLAAGSSLIWAALTVA 85
Db 1 MVSYLVLSCGIVSFQVGFTEPQLILTSVAVLYSGSQFLAGLYGAGASLSVSIITLAF 60
Qy 86 MDRHVLVYGLSLRSRIIQRLOKSKTALWAF----GLTDEVFAAATAKLVRNNRSENW 141
Db 61 LGLRFLVMSSSSRHV-----RQKTTWDFPFSSMTISDESFGVNTVMFSQPD--WTADHA 113

```

```

Qy 142 IGIAFSSWSWVFTVIGAFSGGLQGYPAVEAALGFMLPALFMSFLLASFOKQSLCV 201
Db 114 LALNLLNYGIWVLGSGAL---LVSVDLDTLSIISYGLTAMFICTMTVQFVDY----165
Qy 202 TAALVAGALAGVTLFSPVAIL---AGIVCGCLTALIOAP 237
Db 166 --YLYAGLSVFTIATLVLQNSLGIWVGLLASLIGF 202

RESULT 12
US-10-617-320-3341
; Sequence 3341, Application US/10617320
; Publication No. US20050136404A1
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/617,320
; FILING DATE: 10-Jul-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3341:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...185
; SEQUENCE DESCRIPTION: SEQ ID NO: 3341:
US-10-617-320-3341

Query Match      13.7%; Score 171; DB 5; Length 185;
Best Local Similarity 28.0%; Pred. No. 7.3e-09;
Matches 53; Conservative 38; Mismatches 76; Indels 22; Gaps 9;

Qy 17 EGCDSLPVISYIPVAFAPAGLNATRLG---FSPLESVPFSCIYAGASQFVTMTMLAAG 73
Db 1 EGAQAAMPALGYVSGIACG---TIGAPYVTPVEVGLMSLFYAGSAQFAMALITVQ 56
Qy 74 SSLWIAALTVMAMDMVRHVLGSPSLRSRIIQRLOKSKTALW----AFGLTDEVFAAATAK 128

```

Db 57 APVAIAIATVFLINLR--LFILSLHASTYFR----HTSLWYNIGMSSILTDITGVLMGE 110
Qy 129 LVRNRRNSNMIGIAFSSSSWVFGTVIGAFSGSLGQYPAVEAALGFMLPALFMSF 188
Db 111 LAHTD-KVNPWMHGNLNSVAMFVGTGVTALG-GLLPN-PBI-FGLDPAALVGMFIGI 166
Qy 189 LLASFQKQ 197
Db 167 FASQFQMQ 175

RESULT 13

US-10-156-761-10035
; Sequence 10035, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10035
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10035

Query Match 12.3%; Score 153.5; DB 4; Length 302;
Best Local Similarity 27.8%; Pred. No. 7.8e-07;
Matches 59; Conservative 36; Mismatches 96; Indels 21; Gaps 8;
Qy 20 KDSLPIVISYIPVAFAGNATRLGFSPLSFVFCIIYAGASQFVTAMLAAGSSLUWIA 79
Db 72 RDALGVGVAVGLSGFAGFVTSAGSLSLLTQTCALSLLVFTGASQFALVGLAAGNPLAA 131
Qy 80 ALTVMAMDVHVLVGPISRSRI-IQRLQSKTALMAFGLTDEVFAAATAKLVRNRRNSE 138
Db 132 AAGAFFLGVNRFYGLRLSQLLALPRAVRPFAAQW---VIDETTAVALAQPTRRSVR--- 185
Qy 139 NWMIGI---AFSSSSWVFGTVIGAFSGSLGQYPAVEA-ALGFMLPALFMSFL--LAS 192
Db 186 ---IGFTVTGLSLYVLNLTLLIGAVGA---QAIGDTDANGLDAAAGPAFLALLAPMLK 238
Qy 193 FQKQSLCVTAALVG-ALAGVTLSIPVAILA 223
Db 239 TAARAVAGIAVLGLGLLPLVPAGVPVLVAA 270

RESULT 14

US-09-738-626-4887
; Sequence 4887, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO

; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4887
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4887

Query Match 8.4%; Score 105; DB 3; Length 324;
Best Local Similarity 27.4%; Pred. No. 0.066;
Matches 48; Conservative 26; Mismatches 75; Indels 26; Gaps 8;
Qy 69 MLAGSSSLWIAALTV-NAMDVHVLVGPISRSRI-IQRLQSKTALMAFGLTDEVFAA--- 124
Db 21 MLNNGALVGLIALCVGLFIATPHELTIPNL---INIGIQSATVALLAFGTFVITAGID 77
Qy 125 -ATAKLVRNRRNSNMIGIAFSSSSWVFGTVIGAFSGSLGQYPAVEAALGFMLPA 183
Db 78 LSVGSVAALGAMTSAYFFAEVGLPGWITLLIGLFIGLLAGA-----ISGISIAYG-KLPA 131
Qy 184 LFMSFLASQKQSLCV-----TAALVGLACAGVTLSIPVAILA-----LAGIYC 227
Db 132 FIATLAWMSIARGITLVISQSPISPAVNAL-GRTYFGIPMPILMMALAGIYC 185

RESULT 15

US-09-738-626-6653
; Sequence 6653, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6653
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6653

Query Match 8.2%; Score 102.5; DB 3; Length 474;
Best Local Similarity 22.9%; Pred. No. 0.19;
Matches 58; Conservative 35; Mismatches 87; Indels 73; Gaps 12;

```

Qy 20 KDSLPIVISYIPVA-----FAFGLNATRLGSPLESVPFSCIIYAGASQFVITA 68
Db 60 QDGVLLALGLFLPSWIVRTMLLVAGFAGAWGMBLGPBK-----FLAVTVAIYNPFVVER 114
Qy 69 MLAGSSSL-----WIAALTYMAMDVHVLYGFSLSRIIQRLOKSKTALWAFGL--TDEVF 122
Db 115 LLOGHWSLVMAVWLLPLVVALR--RHP-----RWQVV-----AIWAASLTPTGAVV 158
Qy 123 AAATAKLVRNNRWSNNMIGIAFSSWSWV-----EGTVIGA 160
Db 159 AAVTGVASSKKRFTTL-----CSFLSWLPWLIIPALLATPTSGGALTFAIRSETYAGTLGT 214
Qy 161 PSGSGLLQGYPAVEAA--LGFMLPA-LFMSFLASFORKQSLCVTAALVGALAGV----- 212
Db 215 ALGLGGIWNAGAVPASRELGFPAVAGILLFALLAGFKNCPPWLLALLAVVGFMGAIQPLWM 274
Qy 213 -TLESIPVAILAG 224
Db 275 PNLFTWTIAYVPG 287

```

Search completed: February 15, 2006, 12:29:33
Job time : 158.598 secs

This Page Blank (uspto)

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February.15, 2006, 12:25:52 ; Search time 14.4522 Seconds
(without alignments)
240.922 Million cell updates/sec

Title: US-10-073-293A-4

Perfect score: 1248

Sequence: 1 MESTPQAPGSAFMFGCK.....VCGCLTALIAQFMQAPDEL 245

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 107799 seqs, 14211699 residues

Total number of hits satisfying chosen parameters: 107799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Sorted. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	229.5	18.4	230	6	US-10-793-626-268
2	105	8.4	304	7	US-11-082-389-14
3	91.5	7.3	432	6	US-10-873-528-21
4	91.5	7.3	648	6	US-10-793-626-1060
5	88	7.1	251	7	US-11-054-515-1316
6	87	7.0	472	6	US-10-467-557-2268
7	87	7.0	473	6	US-10-467-557-1874
8	86	6.9	300	7	US-11-098-686-10583
9	86	6.9	469	7	US-11-195-739-19
10	85.5	6.9	453	6	US-10-793-626-3260
11	85.5	6.9	500	7	US-11-012-668-4
12	83.5	6.7	508	6	US-10-467-557-7826
13	83	6.7	539	6	US-10-467-557-3146
14	82.5	6.6	475	7	US-11-074-176-336
15	82.5	6.6	488	7	US-11-074-176-126
16	82	6.6	234	7	US-11-072-512-2791
17	82	6.6	693	6	US-10-467-557-6176
18	81.5	6.5	526	7	US-11-094-317-43
19	81.5	6.5	539	7	US-11-210-316-26
20	81.5	6.5	724	7	US-11-043-889-8
21	80.5	6.5	252	7	US-11-054-515-1326
22	80	6.4	300	6	US-10-858-730-109
23	80	6.4	455	7	US-11-098-686-10673
24	79.5	6.4	463	6	US-10-467-557-6352
25	79.5	6.4	463	6	US-10-467-557-7604

26	79.5	6.4	573	6	US-10-055-877-79	Sequence 79, Appl
27	79.5	6.4	634	7	US-11-169-041-189	Sequence 189, Appl
28	78.5	6.3	248	7	US-11-054-515-1360	Sequence 1360, Ap
29	78.5	6.3	254	7	US-11-054-515-844	Sequence 844, App
30	78.5	6.3	429	7	US-11-205-109-32	Sequence 32, Appl
31	78.5	6.3	439	7	US-11-082-389-80	Sequence 80, Appl
32	78.5	6.3	564	7	US-11-082-389-78	Sequence 78, Appl
33	78.5	6.3	1275	6	US-10-724-598-49	Sequence 49, Appl
34	78	6.2	529	6	US-10-858-730-104	Sequence 104, App
35	78	6.2	529	6	US-10-858-730-105	Sequence 105, App
36	78	6.2	984	7	US-11-055-822-508	Sequence 508, App
37	78	6.2	984	7	US-11-055-822-594	Sequence 594, App
38	77.5	6.2	249	7	US-11-054-515-1321	Sequence 1321, Ap
39	77.5	6.2	650	6	US-10-873-528-95	Sequence 95, Appl
40	77	6.2	249	7	US-11-054-515-957	Sequence 957, App
41	76.5	6.1	316	7	US-11-072-512-2079	Sequence 2079, Ap
42	76.5	6.1	489	6	US-10-858-730-198	Sequence 198, App
43	76.5	6.1	489	7	US-11-055-822-1152	Sequence 1152, Ap
44	76	6.1	448	6	US-10-873-528-42	Sequence 42, Appl
45	75.5	6.0	254	7	US-11-054-515-1659	Sequence 1659, Ap

ALIGNMENTS

RESULT 1
US-10-793-626-268
; Sequence 268, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348008
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 268
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-268

Query Match	18.4%	Score	229.5;	DB	6;	Length	230;
Best Local Similarity	28.6%;	Pred. No.	2.8e-13;				
Matches	65;	Conservative	47;	Mismatches	102;	Indels	13;
Gaps	6;						
Qy	14	TFMEGCKDSLPIVISIYPVAFAGLNATRLGFSVPFSCIIYAGASQFVITAMLAAG	73				
Db	5	TFQGVKCEIPTLLGAGVGLSFGIVAVSQNFSLVLEITLLCLIIYAGAAQFIICTLVIAG	64				
Qy	74	SSLIWALATVMADMVHRVLYGPSLRRIIOQLQSKTALWAFGLTDFVFAAATAKLVRNN	133				
Db	65	TPISAIVTLTILVNSRMFLSLMTLAPNYKYQGFNRVGLGTL-LTDETFGVAITPYK-G	122				
Qy	134	RRNSNMIGIAFSSSSSWVFGTVIGAFSGSLGQYPAVEAALG--FMLPALFMSFLA	191				
Db	123	EKINDRWLHGLNITAVLFTWVSCVIGAFGE----YISNPDALGLDFAITAMFICLS	177				
Qy	192	SFQ--RKQSLCVTAALVAGLAGVTLF--SIPVAILAGIVCGCLTALI	234				
Db	178	QFEGIKKRLRIYIVLVICVIVMMLLSILPSYLAIIAIAVAALL	224				
RESULT 2							
US-11-082-389-14							
; Sequence 14, Application US/11082389							
; Publication No. US20050244935A1							

REFUGIARY: FUGA Mollayazza

Qy	124	AATAKLVRNRRWSENWNMIGIAFSS-----WSGVVGFOTV-----	157
		: :	
Db	238	KKYAGLEQLG-----VYSWGISFGGAALLQSIFSTVTVPYIPRAIBENATPARLSATAE	292
		: :	
Qy	158	-----IGAFS-----GGLLL-QGYPAVE-AALGFMLPALP-----MSFLLASFORK	196
		: :	
Db	293	SAAAILASALCLTGTFSPSLAIIIPENYAARFTVWSCMLPPLPYTLTETSGIGLNVWRK	352
		: :	
Qy	197	QSLCVTAALVGALAGVTTFSPVAILAGIVGCCTALIQAFW	238
		: :	
Db	353	-TRPIALATLGALA-ANILLICLAVPSGCTRGAAVAACAASF	392
		: :	

```

RESULT 8
US-11-098-686-10583
; Sequence 10583, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FASTSEQ for Windows version 4.0
; SEQ ID NO 10583
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10583

```

[illegible]

RESULT 9
US-11-195-739-19
; Sequence 19, Application US/11195739
; Publication No. US20060019324A1
; GENERAL INFORMATION:
; APPLICANT: ALZARI, PEDRO
; APPLICANT: BOITEL, BIRGITTE
; APPLICANT: VILLARINO, ANDREA
; APPLICANT: FERNANDEZ, PABLO
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: PKNB KINASE AND PSTP PHOSPHATASE AND METHODS OF IDENTIFYING INHIBITORY SUBSTANCES
; TITLE OF INVENTION: INHIBITORY SUBSTANCES

```

; FILE REFERENCE: 252853US
; CURRENT APPLICATION NUMBER: US/11/195,739
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US/10/892,170
; PRIOR FILING DATE: 2004-07-16
; PRIOR APPLICATION NUMBER: US 60/487,943
; PRIOR FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 19
; LENGTH: 469
; TYPE: prt
; ORGANISM: Mycobacterium tuberculosis
; US-11-195-739-19

```

Query Match	6.98; Score 86; DB 7; Length 469;
Best Local Similarity	23.08; Pred. No. 2;
Matches	51; Conservative 32; Mismatches 83; Indels 56; Gaps 11;
Qy	44 GFS-PLE-----SVFFSCIIYAGASQFVITAMLAAGSSL-----WIAALTV 83
Db	180 GFIOIQAEFSKILLIIPFSVAIVAKRLFTSAGKHLIGMTLPRPRDLAPLLAAWVISGVG 239
Qy	84 MAMD-----VRHVLYGPSLRSRIIORLOKSKTALWAFGLTDEVFAAATAK---LVNRNRW 136
Db	240 MVFEKOLGASLLLYTSL---VVVYLATORFSWVVIGLT--LFAAGTLVAYFIFEHVRLR 294
Qy	137 SENNMIGIAFSSWSW-----VFGTVIGAFSGSGLLQGYPAVEAALGFMLPALFMFSLLA 191
Db	295 VQTWLDFDPPDGGYQIVQSLFSFATGGTFFGLGNGQPD-----VFPAASTDFIIA 347
Qy	192 SFKOSKSLCVTAALVAGALAGVTILFSIPVAILAGIVCGCLTAL 233
Db	348 AFREE-----IGHVLGTAILMLTYIV-----IIRGLRTAI 377

```

RESULT 10
US-10-793-626-3260
; Sequence 3260, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 3260
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3260

```

```

Query Match      6.9%; Score 85.5; DB 6; Length 453;
Best Local Similarity 23.4%; Pred. No. 2.1;
Matches 57; Conservative 42; Mismatches 92; Indels 53; Gaps 12;

Qy 23 LPVISVIPAF--AFGLNATRIGFSPLESVPFSCI-----IYAGASQFVITAML 70
      :::: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 146 IPLVGALITVDFVLLFIIMKFGFKRIEAVGLIFTVLAIAPFVEVVISPPQ--ITDML 203
      : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 71 -----AAGSSLIAALTVMADVRLVYGPSLRSRRIIQLQKSKTALWAFGLTDE 120
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 NGFVPKHEIITNQGIYALIGAITMPHNL---LHSSIVQSKYDRHSI-----HE 254
      : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 121 VFRAATAKLVRRNRRRWSENNMIG-IAPSSWSSWVFGTV---IGAFSSGGL-LQGYPAVEA 175
      : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 255 KAQAIVATDSNTQLSIAFVFNCLLITLGAALFFGTTEDLGGFYDLYLALKTEPALGA 314
 Qy 176 ALGMLPALFMSFLLASFORQSLCVTAALVG--ALAGVTLFSIP-----VAILA 223
 Db 315 TLGGIMSTLFAVALLAS---QNSTITGTLAGQIVMEGFLKLSIPNWLRLITRSLAIVP 371
 Qy 224 GIVC 227
 Db 372 VIIC 375

RESULT 11
 US-11-012-668-4
 ; Sequence 4, Application US/11012668
 ; Publication No. US20050010512A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Stephen M.
 ; APPLICANT: Rafalski, J. Antoni
 ; TITLE OF INVENTION: Nitrogen Transport Metabolism
 ; FILE REFERENCE: BB-1210
 ; CURRENT APPLICATION NUMBER: US/11/012.668
 ; CURRENT FILING DATE: 2004-12-15
 ; PRIOR APPLICATION NUMBER: 60/098,248
 ; PRIOR FILING DATE: 28 August 1998
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 4
 ; LENGTH: 500
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 US-11-012-668-4

Query Match 6.9%; Score 85.5; DB 7; Length 500;
 Best Local Similarity 23.1%; Pred. No. 2.4;
 Matches 65; Conservative 30; Mismatches 89; Indels 97; Gaps 16;

Qy 29 YIPVAFAPGLNATRLGP-----SPLESVFSCIIYAGASQFVITAM-LAAGSLW 77
 Db 97 YLFGFAPAFGSPNGPIGKHFFGLKDIPISSYDYSYFLYQWA--FAIAAAGITSGS-- 151
 Qy 78 IALITWAMDVRHVLGSPRSRIIQLQSKTALW-----AFGLTDEVFPA-- 124
 Db 152 IAERTQP---VAYLISSFLTGFPYVVS---NFWSPDGNWASAFKITDRLFTSGVID 203
 Qy 125 -ATAKLVRNRRNRSNWMIGIAFSSWVFGTVIGAFSGG--LLOGYPAVEAALG-F 179
 Db 204 FAGSGVH-----WVGIIAGLWGLIEGPRMGRFDHAGRAVALRGHSLSVLVLTGF 254
 Qy 180 ML-----PALFMSFL-----ASFQKQSLCVTAALVGLAGV-TLFSIPV-- 219
 Db 255 LLWFGWYGFNGFSFNKILLYTNGSNGYNGYQWSAVGRVAVTTTLAGSTAALTTLFGKRVIS 314
 Qy 220 -----AIIAG-----IVCGCLTALI 234
 Db 315 GHNVTVDVNCNLLGGLGFAAITAGCSVVEPMAAIVCGFVASIV 355

RESULT 12
 US-10-467-657-7826
 ; Sequence 7826, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SpA
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MASIGNANI Vega
 ; APPLICANT: MONACI Elisabetta
 ; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/467,657
 ; CURRENT FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: GB-0103424.8
 ; PRIOR FILING DATE: 2001-02-12

; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: SeqWin99, version 1.04
 ; SEQ ID NO 7826
 ; LENGTH: 508
 ; TYPE: PRT
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-7826

Query Match 6.7%; Score 83.5; DB 6; Length 508;
 Best Local Similarity 18.9%; Pred. No. 3.6;
 Matches 63; Conservative 39; Mismatches 108; Indels 123; Gaps 11;


Qy 25 IVIYIIPVAFAPGLNATRLGFSPL-ESVFFSCIIYAGASQFVITAMLA--G 73
 Db 129 IILFFFTIYCASGIVAGATLFSQSLFEGMTYNQAMWLGAGATIAVTLGGFLAVSWTDTLQ 188
 Qy 74 SSLWIAALTVMAMDVRHVLGSPRSRIIQLQSKS-----KTALWAF 115
 Db 189 ASLMIFALITLTPVMVYVGLGGABQMSAAIQSVAAGTKGYSLFAGTTVIGIITAAWGL 248
 Qy 116 G-----LTDEVFPAATAKLVRNRRNRSNWM-----IGIAF-----SSW 149
 Db 249 GYFGOPHILARFMAESAKSLVSARRIGMTWMAALCLAGAVAGVFGIAYFGANPDKVS 308
 Qy 150 SS-----WVFGTVIGAFSGSLGQYPAVEAALGFM-----PALF 185
 Db 309 SGNHERIFIALSTLLFPNPIAGIILSAI-----LAAVMSTLSCQLLVCSAITEDFY 360
 Qy 186 MSLLASFORQSLCVTAALVGLAGVTLFSI-----PVAI-- 221
 Db 361 KGFLRKNAQOQSELVWVGRMLVLAIVISILIASDPNSKVLGLVSYAWAGFAAGFPIVIL 420
 Qy 222 -----LAGIVCGCLTALIOAFWOGAP 242
 Db 421 SVLWKRITAYGALSGNVAGASTVVMWAEWVKP 453

RESULT 13
 US-10-467-657-3146
 ; Sequence 3146, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SpA
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MASIGNANI Vega
 ; APPLICANT: MONACI Elisabetta
 ; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/467,657
 ; CURRENT FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: GB-0103424.8
 ; PRIOR FILING DATE: 2001-02-12
 ; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: SeqWin99, version 1.04
 ; SEQ ID NO 3146
 ; LENGTH: 539
 ; TYPE: PRT
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-3146

Query Match 6.7%; Score 83; DB 6; Length 539;
 Best Local Similarity 22.1%; Pred. No. 4.3;
 Matches 60; Conservative 41; Mismatches 100; Indels 70; Gaps 15;

Qy 9 APGSATFMGCKDLSLPIVSIYIPVAF--AFGLNATRLGSPLESVFSCIIYAGASQFVI 66
 Db 240 AMGLSHLDREKHTYVFLLLGIAYCASIGGLGLTVGSGPP-----NLIAAKALNLD 292
 Qy 67 TAWLAAGSSIIWIAALTVMAMDVRHVLGSPRSRIIQLQSK-----SKTA 111
 Db 293 VGMKLGLPMMILLIPLMLLSL--YVILKPNLNERVBIKAESIPWTLHRVALLIFLATA 351

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using  model

Run on: February 16, 2006, 17:05:59 ; Search time 6212.4 Seconds
(without alignments)
2241.749 Million cell updates/sec

Title: US-10-073-293A-4
Perfect score: 1248
Sequence: 1 MESPTQPAPGSAATFMEGCK.....VCGCLTALIQAFWQAPDEL 245

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 20421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/absz/ABSSWEB spool/US10073293/runat_15022006_120734_963/app_query.fasta.1
-DB=GenEmbl -QWMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=absz04
-USER=US10073293 @CGN 1.1 5142 @runat_15022006_120734_963 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREAU=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_hng:*
15: gb_pi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1248	100.0	738	6	BD177948 Process f
2	1248	100.0	738	6	CS078097 Sequence
3	1248	100.0	738	6	CS078113 Sequence

4	1248	100.0	738	6	AX534668
5	1248	100.0	19150	1	D90891
6	1248	100.0	110000	1	U00096_27
7	1248	100.0	110000	1	U00096_28
8	1247	99.9	110000	1	AE005174_35
9	1247	99.9	110000	1	AE005174_36
10	1247	99.9	110000	1	BA000007_35
11	1247	99.9	305325	1	AE016765
12	1243	99.6	110000	1	AE005674_27
13	1243	99.6	290380	1	AE016987
14	1013.5	81.2	756	6	AR386287
15	871.5	69.8	10733	1	AE013695
16	871.5	69.8	220050	1	AJ414156
17	871.5	69.8	290002	1	AE017129
18	871.5	69.8	110000	1	EX936398_10
19	845.5	67.7	110000	1	EX950851_39
20	841.5	67.4	243072	6	AX770904
21	841.5	67.4	349318	1	EX571863
22	834	66.8	3619	1	EC410307
23	592	47.4	801	6	AR319849
24	431	34.5	190822	14	AC051613
25	385	30.8	348624	1	EX640441
26	385	30.8	348666	1	EX640426
27	362	29.0	343243	1	EX640414
28	347	27.8	202872	14	AC016160
29	317.5	25.4	13099	1	AE000982
30	308.5	24.7	6002	1	AF118548
31	308.5	24.7	110000	1	AE014291_17
32	308.5	24.7	110000	1	AE017223_17
33	308.5	24.7	110000	1	AE017223_18
34	301.5	24.2	349648	1	EX957220
35	300.5	24.1	110000	1	AP008226_02
36	300.5	24.1	247910	1	AE017307
37	287	23.0	110000	1	BA000004_30
38	281	22.5	110000	1	BA000043_06
39	278.5	22.3	110000	1	AP006840_04
40	273	21.9	110000	1	CP000900_05
41	269	21.6	110000	1	AE017225_16
42	269	21.6	110000	1	AE017334_16
43	269	21.6	110000	1	AE017355_16
44	269	21.6	290525	1	AE017029
45	268.5	21.5	110000	1	AE017333_10

ALIGNMENTS

RESULT 1	BD177948	Process for producing L-amino acid using escherichia.	738 bp	DNA	linear	PAT 16-APR-2003
LOCUS	BD177948	Process for producing L-amino acid using escherichia.				
DEFINITION	BD177948	Process for producing L-amino acid using escherichia.				
ACCESSION	BD177948	Process for producing L-amino acid using escherichia.				
VERSION	BD177948.1	GI:30015211				
KEYWORDS	JP 2002300874-A/3.	Escherichia coli				
SOURCE	Escherichia coli	Escherichia coli				
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.					
REFERENCE	1 (bases 1 to 738)					
AUTHORS	Tabolina,E.A., Rybak,K.V., Khourges,E.M., Voroshilova,E.B. and Gulyatiner,M.M.					
TITLE	Process for producing L-amino acid using escherichia					
JOURNAL	Patent: JP 2002300874-A 3 15-OCT-2002;					
COMMENT	AJINOMOTO CO INC					
	OS Escherichia coli					
	PN JP 2002300874-A/3					
	PD 15-OCT-2002					
	PR 13-FEB-2002 JP 2002034760					
	PR 13-FEB-2001 RU 2001103865, 26-FEB-2001 RU 2001104998 PR					
	26-FEB-2001 RU 2001104999, 28-JUN-2001 RU 2001117632 PR					
	28-JUN-2001 RU 2001117633					
	PI EKATERINA ALEKSANDROVA TABOLINA, KONSTANTIN VYACHESLAVOVICH					
	PI RYBAK,					
	PI EVGENI MOISEVICH KHOURGES, ELVIRA BORISOVNA VOROSHILOVA, PI					

```

MIKHAIL MAROVICH GUSYATINER
PC C12N1/21,C12P13/04,C12P13/06,C12P13/08,C12P13/12, PC
C12P13/24//
PC (C12N1/21,C12R1:19),(C12P13/04,C12R1:19),(C12P13/06,C12R1:19),
PC (C12P13/08,C12R1:19),(C12P13/08,C12R1:19),(C12P13/12,C12R1:19), PC
(C12P13/24,C12R1:19)
CC Process for producing L-amino acid using escherichia FH Key
FT CDS Location/Qualifiers
FEATURES
    source      1..738
                /organism="Escherichia coli"
                /mol_type="genomic DNA"
                /db_xref="taxon:562"
ORIGIN
Alignment Scores:
Pred. No.:      5.54e-107      Length:      738
Score:          1248.00      Matches:      245
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             6           Gaps:         0
US-10-073-293A-4 (1-245) x BD177948 (1-738)
Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db 1 ATGGAAGCCCTACTCCACAGCCTGCTCCTGGTTCGGCGACCTTCATGGAAGGATGCAAA 60
Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
Db 61 GACAGTTTACCAGATTGTTATTAGTTATATTCCGGTGGCCCTTTCGGTCTCGAATGCG 120
Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
Db 121 ACCCGTCTGGGATTCCTCTCTCGAAGCGTTTTTTTCTCTGCATCATTTATGCGAGC 180
Qy 61 AlaSerGlnPheValIleThrAlaMetLeuAlaLaGlySerSerLeuTrpIleAlaAla 80
Db 181 GCGAGCCAGTTTCGTCAATACCGCATGCTGGCAGCGCGGAGTAGTTTGTGATTCGTGCA 240
Qy 81 LeuThrValMetAlaMetAspValArgHisValIleuTyrGlyProSerLeuArgSerArg 100
Db 241 CTGACCGTCATGGCAATGGATGTTCCGCCATGTGTTGTTATGCGCCGTCCTGCGTAGCCGT 300
Qy 101 IleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu 120
Db 301 ATTATTACGCGTCTGCAAAATTCGAAACCGCCCTGTGGGCTTGGGCTGACGGATGAG 360
Qy 121 ValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTTPSerGluAsnTrp 140
Db 361 GTTTTTCGCGCGCAACCGCAAACTGGTAGCAATAATCGCCGCTGGAGCGAGAACTGG 420
Qy 141 MetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIleGlyAla 160
Db 421 ATGATCGCGATGCGCTTCAGTTCATGGTCATCGTGGGTATTGTTGGTACGGTAATAGGGCA 480
Qy 161 PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet 180
Db 481 TTCTCCGCGAGCGGCTTCGTCGAAGTTATCCCGCGTTGAGCTGCATTAGTTTATG 540
Qy 181 LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys 200
Db 541 CTTCCGGCACCTCTTTATGAGTTTCCTGCTCGCTCTTCCAGCGCAACAATCTCTTTCG 600
Qy 201 ValThrAlaAlaLeuValIleAlaLeuAlaGlyValThrLeuPheSerIleProValAla 220
Db 601 GTTACCGCAGCGGTAGTTAGTTGGTCCCTTGCAGCGGTAACGCTATTTCCTATCCCGTCGCC 660
Qy 221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly 240

```

```

Db 661 ATTCTGCGAGGCAATTGCTGTGGTGGCTGCTCACTCGTTAATCCAGGCAATTCGTGCAAGGA 720
Qy 241 AlaProAspGluLeu 245
Db 721 GCGCCCGATGAGCTA 735
RESULT 2
CS078097 738 bp DNA linear PAT 06-MAY-2005
LOCUS Sequence 3 from Patent EP1526179.
DEFINITION CS078097
ACCESSION CS078097
VERSION CS078097.1 GI:63093065
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 Tabolina, E.A., Rybak, K.V., Khourges, E.M., Voroshilova, E.B. and
Gusyatiner, M.M.
TITLE Method for producing L-amino acid using bacteria belonging to the
Genus Escherichia
JOURNAL Patent: EP 1526179-A 3 27-APR-2005;
Ajinomoto Co., Inc. (JP)
FEATURES
    Location/Qualifiers
    source      1..738
                /organism="Escherichia coli"
                /mol_type="unassigned DNA"
                /db_xref="taxon:562"
    CDS          1..738
                /note="unnamed protein product"
                /codon_start=1
                /transl_table=11
                /protein_id="CA194223.1"
                /db_xref="GI:63093066"
                /translation="MAGSTPQAPFAGSATFMEGCKDSLPIVISIYPVAFAGLNATRLG
                FSPLESYFFSCIIYVAGASQFIVTAAAGSSLTAAITVMAMDVVRHLYGSLRSRII
                ORLQKSTALWAFGLTDEVFAAATAKLVRNNRRWSNNMIGIAFSSWSWVGTVIGA
                FSGSLGLQGYPAVEALGFMPLPALFMSFLLASFORKQSLCVTAALVGLAGLVTFISIP
                VALLAGIVCGCLTALIQAFWQGAPEL"
    ORIGIN
Alignment Scores:
Pred. No.:      5.54e-107      Length:      738
Score:          1248.00      Matches:      245
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             6           Gaps:         0
US-10-073-293A-4 (1-245) x CS078097 (1-738)
Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db 1 ATGGAAGCCCTACTCCACAGCCTGCTCCTGGTTCGGCGACCTTCATGGAAGGATGCAAA 60
Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
Db 61 GACAGTTTACCAGATTGTTATTAGTTATATTCCGGTGGCCCTTTCGGTCTCGAATGCG 120
Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
Db 121 ACCCGTCTGGGATTCCTCTCTCGAAGCGTTTTTTTCTCTGCATCATTTATGCGAGC 180
Qy 61 AlaSerGlnPheValIleThrAlaMetLeuAlaLaGlySerSerLeuTrpIleAlaAla 80
Db 181 GCGAGCCAGTTTCGTCAATACCGCATGCTGGCAGCGCGGAGTAGTTTGTGATTCGTGCA 240
Qy 81 LeuThrValMetAlaMetAspValArgHisValIleuTyrGlyProSerLeuArgSerArg 100
Db 241 CTGACCGTCATGGCAATGGATGTTCCGCCATGTGTTGTTATGCGCCGTCCTGCGTAGCCGT 300
Qy 101 IleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu 120

```


Db 301 ATTATTCAGCTCTGCAAAATCGAAAAACCCGCTGTGGGCTTTGGCTGACGGATGAG 360
121 ValPheAlaAlaAlaThrAlaLysLeuValArgAenAenArgATpSerGluAenTtp 140
361 GTTTTGGCGCGCAACCGCAAACTGGTAGCAATATATCCGCTGGAGCGAGACTGG 420
141 MetIleGlyIleAlaPheSerSerTpsSerTtpValPheGlyThrValIleGlyAla 160
421 ATGATCGGCATTCGCTTCAGTTTCATGGTTCATCGTGGGTATTGGTACGGTAATAGGCGCA 480
161 PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet 180
481 TTCTCCGCGACGCGCTTGTGCAAGGTATATCCCGCTGAAGCTGCAATAGGTTTTATG 540
181 LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnAlaSerGlnLysGlnSerLeuCys 200
541 CTTCGCGCACTCTTTATGAGTTTCTGCTCGCCTCTTTCCAGCGCAACAATCTCTTTGGC 600
201 ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla 220
601 GTTACCGCAGCGTTAGTTGGTGCCCTTGCAGGCGTAACGCTATTCTATTCCGCTCGCC 660
221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTtpGlnGly 240
661 ATCTGGCAGGCATTTGTGTGGCTGCTCATCTCGTTAATCCAGGCATTTCTGGCAAGGA 720
241 AlaProAspGluLeu 245
721 GCGCCCGATGAGCTA 735

RESULT 3
CS078113
LOCUS CS078113 738 bp DNA linear PAT 06-MAY-2005
DEFINITION Sequence 3 from Patent EP1526181.
ACCESSION CS078113
VERSION CS078113.1 GI:63093081
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Escherichia coli

REFERENCE 1
AUTHORS Tabolina,E.A., Rybak,K.V., Khourges,E.M., Voroshilova,E.B. and Gussyatiner,M.M.
TITLE Method for producing l-amino acid using bacteria belonging to the Genus Escherichia
JOURNAL Patent: EP 1526181-A 3 27-APR-2005;
Ajinomoto Co., Inc. (JP)
FEATURES
source 1. 738
location/Qualifiers
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
1. 738
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAI94227.1"
/translation="MESPTQPAPFGSATFMGCKDSLPIVSIYPVAFGLNATRLG
FSPLESVFFSCIIIVAGASQFVITAMLAAGSLMTAALTVMAMDVHLYGFSLSRII
QRLOKSTALWAFGLTDEVFAAATAKLVRNNRWSNNMIGIAFSSWSVVFQVIGA
FSGGLLQGPVPAEVALGFMLPALFMSFLASFORQSLCVTAALVAGALAGVTLFSP
VAILAGVCGCLTALIQAFWQGFDEL"

ALIGNMENT SCORES:
Pred No.: 5 54e-107 Length: 738
Score: 1248.00 Matches: 245
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0

ORIGIN
Alignment Scores:
Pred No.: 5 54e-107 Length: 738
Score: 1248.00 Matches: 245
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0

DB: 6 Gaps: 0
US-10-073-293A-4 (1-245) x CS078113 (1-738)
Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db 1 ATGAAAGCCCTACTCCACAGCCTGCTCTGGTTTCGGCGACCTTCATGGAGGATGCAAA 60
Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
Db 61 GACAGTTTACCGATTGTTATTAGTTATATTCGGTGGCCTTTGGTTTCGGTCTGAATGCG 120
Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
Db 121 ACCCGTCTGGGATTCCTCCCTCTCGAAAGCGCTTTTTCCTCTCGCATCATTTATGCGAGC 180
Qy 61 AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTtpIleAlaAla 80
Db 181 GCGAGCCAGTTTCGTCATTACCGGATGCTGGCAGCGGAGTAGTTTGTGATTTGCTGCA 240
Qy 81 LeuThrValMetAlaMetAspValArgHisValIleuTyrGlyProSerLeuArgSerArg 100
Db 241 CTGACCGCTCATGGCAATGGATGTTCCGCATGTGTGTATGCCCGCTCCTGCGTAGCCGT 300
Qy 101 IleIleGlnArgLeuGlnLysSerLysThrAlaLeuTtpAlaPheGlyLeuThrAspGlu 120
Db 301 ATTATTTCAGCGCTCTGCAAAATCGAAACCGCCCTGTGGCGGTTTGGCCTGACGATGAG 360
Qy 121 ValPheAlaAlaThrAlaLysLeuValArgAenAenArgTtpSerGluAenTtp 140
Db 361 GTTTTGGCGCGCAACCGCAAACTGGTAGCAATATATCCGCTGGAGCGAGACTGG 420
Qy 141 MetIleGlyIleAlaPheSerSerTpsSerTtpValPheGlyThrValIleGlyAla 160
Db 421 ATGATCGGCATTCGCTTCAGTTTCATGGTTCATGTGGTATTGGTAGCGTAATAGGCGCA 480
Qy 161 PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet 180
Db 481 TTCTCCGCGACGCGCTTGTGCAAGGTATATCCCGCTGTTCAAGCTGATAGGTTTTATG 540
Qy 181 LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnLysGlnSerLeuCys 200
Db 541 CTTCGCGCACTCTTTATGAGTTTCTGCTCGCCTCTTTCCAGCGCAACAATCTCTTTGGC 600
Qy 201 ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla 220
Db 601 GTTACCGCAGCGTTAGTTGGTGCCCTTGCAGGCGTAACGCTATTCTATTCCGCTCGCC 660
Qy 221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTtpGlnGly 240
Db 661 ATCTGGCAGGCATTTGTGTGGCTGCTCATCTCGTTAATCCAGGCATTTCTGGCAAGGA 720
Qy 241 AlaProAspGluLeu 245
Db 721 GCGCCCGATGAGCTA 735

RESULT 4
AX534668
LOCUS AX534668 738 bp DNA linear PAT 22-NOV-2002
DEFINITION Sequence 3 from Patent EP1239041.
ACCESSION AX534668
VERSION AX534668.1 GI:25261073
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Escherichia coli
REFERENCE 1
AUTHORS Tabolina,E.A., Rybak,K.V., Khourges,E.M., Voroshilova,E.B. and Gussyatiner,M.M.
TITLE Method for producing l-amino acid using bacteria belonging to the Genus escherichia
JOURNAL Patent: EP 1239041-A 3 11-SEP-2002;

E-mail: hmori@gtc.aist-nara.ac.jp
URL:
The Japan E. coli genome database
http:bsw3.aist-nara.ac.jp.

FEATURES

source

1. .19150
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K12"
/db_xref="taxon:562"
/map="60.2 min"
/clone="Kohara clone #445"
/clone_lib="Kohara lambda miniset library"
/note="Nucleotide position 2791307-2810456 from the
initiation site of ThrA (0 min.).
This clone is from Kohara lambda miniset library"
complement(1. .253)
/gene="yzzM"
complement(<1. .>253)
/gene="yzzM"
/note="similar to [SwissProt Accession Number P39169]
start codon is not identified yet"
/codon_start=1
/transl_table=11
/product="UNKNOWN PROTEIN FROM 2D-PAGE (SPOT LM6)
(FRAGMENT)." ."
/protein_id="BAA16532.1"
/db_xref="GI:1800055"
/translation="GLFNFVKDAGEKLWDVATGQHKDDQAKVQEHNTKGIPTADK
VNIQADKATVVGDLGSQSAKEKILVAVNGISIASVDD"
complement(340. .498)
/gene="ZK632.10"
complement(340. .498)
/gene="ZK632.10"
/note="similar to [SwissProt Accession Number P34655]"
/codon_start=1
/transl_table=11
/protein_id="BAA16533.1"
/db_xref="GI:1800056"
/translation="MGFWRVITIIPLPLGLVGKGGWAFIINILLTLGYPGLIH
AFWVQRD"
681. .980
/gene="HLYU"
681. .980
/gene="HLYU"
/note="similar to [SwissProt Accession Number P52695]"
/codon_start=1
/transl_table=11
/product="TRANSCRIPTIONAL ACTIVATOR HLYU."
/protein_id="BAA16534.1"
/db_xref="GI:1800057"
/translation="MTELAQLQASAEQAAALLKAMSHPKELLILCMLSGSPGTSAGEL
TRTGLSASATSOHLARMREGLDSQDQRIILYSIKNEAVNAIATLKNVYCP"
complement(1561. .1965)
/gene="hnsB"
complement(1561. .1965)
/gene="hnsB"
/note="similar to [SwissProt Accession Number P30017]"
/codon_start=1
/transl_table=11
/product="DNA-BINDING PROTEIN STPA."
/protein_id="BAA16535.1"
/db_xref="GI:1800058"
/translation="MSVMQLSNIRTLRAMAREFSIDVLEEMLEKFRVVTKRRREE
EQOQRELAERQEKISTVLELMKADGINPELLGNSSAAAPRAGKKRQPRPKYKFTDV
NGTKTWTGQGRTPKPTAALAEKSLDDFLI"
complement(3120. .3476)
/gene="ygaC"
complement(3120. .>3476)
/gene="ygaC"
/note="similar to [SwissProt Accession Number P36931]
start codon is not identified yet"

gene

CDS

gene
3604. .3945
/gene="yqjD"
3604. .3945
/gene="yqjD"
/note="similar to [SwissProt Accession Number P42617]"
/codon_start=1
/transl_table=11
/protein_id="BAA16537.1"
/db_xref="GI:1800060"
/translation="MGDHPNRPNRNDVDGVQDIONDVNLQADLSLVLSKWSGSDAK
GEAAARSAQAALLKETRAMHGRTRVQQAARDVAGCADSFVRERPWCSTGTAAGA
FIGALLSMRKS"
4193. .4846
/gene="MG230"
4193. .4846
/gene="MG230"
/note="similar to [SwissProt Accession Number P47472]"
/codon_start=1
/transl_table=11
/protein_id="BAA16538.1"
/db_xref="GI:1800061"
/translation="MRITTYRNDCCVCHATKAMENRGDFDFEMINVDVRYPERAEALR
AQGFQPLVVIAGDLSWGSGRPMINRLASSATRGCMQLYVFSSSENTQAFIERL
GLPVRIPLNERIRIQVDEPYILIVPSYGGGTAGVPRQVIRFLNDEHNRALLRGI
ASGRNPFGEAYGRAGDVIARKCGVPLYRFELMQTSQDIENVRKGVTEFWQRPQNA"
4858. .6414
/gene="NRDE"
4858. .6414
/gene="NRDE"
/note="similar to [SwissProt Accession Number Q08698]"
/codon_start=1
/transl_table=11
/product="RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE 2 ALPHA
CHAIN (EC 1.17.4.1) (RIBONUCLEOTIDE REDUCTASE 2)."
/protein_id="BAA16539.1"
/db_xref="GI:1800062"
/translation="MDYHALNMLNLYDSAGRIQFDKDRQAVDAFIATHVRPNSVTF
SQQRNLWLVNDEGYDESIVNRYSRDFTVLTHTAHTSGRFOTFLGAWFYTSYTLK
TDGKYLEDFADRVMTVALTLAQDETALQLTDEMLSGRFQFATFLNCGKQQRG
ELVSCFLRIEDNMESIGRAVNSALQISKRGGVAFLLSNLRAGAPIKKIENSSGV
IPVKMELEDAFSYANQLGARQAGAVILHAHPDILRFLDTKRENADEKIRIKTSLUG
WIPDITFLAKENAOALFSFYDVERVYKGFADVAISQHYDELVADEIRIKYLLNA
RDFQRLAIOFESGYPIYMYEDTVNRPVAGRIINMSNLCSILQVNSASEYDENLD
YRTGHDISCNLSLNIATMDSDFARTVETAVRGLTAVSDMSHRSVPSIAGNAA
SHAIGIGQNLHGYLAREGIAYGSPALDFTNLYFYAITWHALRTSMLARERGETTA
GFKQSRYAQ"
6272. .6964
/note="similar to [PIR Accession Number S34271]"
/codon_start=1
/transl_table=11
/product="ribonucleoside-diphosphate reductase (EC
1.17.4.1)"
/protein_id="BAA16540.1"
/db_xref="GI:1800063"
/translation="MVRRKHWSPISSIMSPGCMHCVPRCCWHANA VKPSPGNSHAM
PSGEYFQYVQLGNWQPKTAKVGLFTRSGITLPTREMAQLRDDVMRYLYNQLQAV
PPTGISYINHTASSIHPVAKVEIRKEGTGVYYPAPPTMTNENLALYQDAEIGE

gene

CDS

gene
4193. .4846
/gene="MG230"
4193. .4846
/gene="MG230"
/note="similar to [SwissProt Accession Number P47472]"
/codon_start=1
/transl_table=11
/protein_id="BAA16538.1"
/db_xref="GI:1800061"
/translation="MRITTYRNDCCVCHATKAMENRGDFDFEMINVDVRYPERAEALR
AQGFQPLVVIAGDLSWGSGRPMINRLASSATRGCMQLYVFSSSENTQAFIERL
GLPVRIPLNERIRIQVDEPYILIVPSYGGGTAGVPRQVIRFLNDEHNRALLRGI
ASGRNPFGEAYGRAGDVIARKCGVPLYRFELMQTSQDIENVRKGVTEFWQRPQNA"
4858. .6414
/gene="NRDE"
4858. .6414
/gene="NRDE"
/note="similar to [SwissProt Accession Number Q08698]"
/codon_start=1
/transl_table=11
/product="RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE 2 ALPHA
CHAIN (EC 1.17.4.1) (RIBONUCLEOTIDE REDUCTASE 2)."
/protein_id="BAA16539.1"
/db_xref="GI:1800062"
/translation="MDYHALNMLNLYDSAGRIQFDKDRQAVDAFIATHVRPNSVTF
SQQRNLWLVNDEGYDESIVNRYSRDFTVLTHTAHTSGRFOTFLGAWFYTSYTLK
TDGKYLEDFADRVMTVALTLAQDETALQLTDEMLSGRFQFATFLNCGKQQRG
ELVSCFLRIEDNMESIGRAVNSALQISKRGGVAFLLSNLRAGAPIKKIENSSGV
IPVKMELEDAFSYANQLGARQAGAVILHAHPDILRFLDTKRENADEKIRIKTSLUG
WIPDITFLAKENAOALFSFYDVERVYKGFADVAISQHYDELVADEIRIKYLLNA
RDFQRLAIOFESGYPIYMYEDTVNRPVAGRIINMSNLCSILQVNSASEYDENLD
YRTGHDISCNLSLNIATMDSDFARTVETAVRGLTAVSDMSHRSVPSIAGNAA
SHAIGIGQNLHGYLAREGIAYGSPALDFTNLYFYAITWHALRTSMLARERGETTA
GFKQSRYAQ"
6272. .6964
/note="similar to [PIR Accession Number S34271]"
/codon_start=1
/transl_table=11
/product="ribonucleoside-diphosphate reductase (EC
1.17.4.1)"
/protein_id="BAA16540.1"
/db_xref="GI:1800063"
/translation="MVRRKHWSPISSIMSPGCMHCVPRCCWHANA VKPSPGNSHAM
PSGEYFQYVQLGNWQPKTAKVGLFTRSGITLPTREMAQLRDDVMRYLYNQLQAV
PPTGISYINHTASSIHPVAKVEIRKEGTGVYYPAPPTMTNENLALYQDAEIGE

gene

CDS

gene
4193. .4846
/gene="MG230"
4193. .4846
/gene="MG230"
/note="similar to [SwissProt Accession Number P47472]"
/codon_start=1
/transl_table=11
/protein_id="BAA16538.1"
/db_xref="GI:1800061"
/translation="MRITTYRNDCCVCHATKAMENRGDFDFEMINVDVRYPERAEALR
AQGFQPLVVIAGDLSWGSGRPMINRLASSATRGCMQLYVFSSSENTQAFIERL
GLPVRIPLNERIRIQVDEPYILIVPSYGGGTAGVPRQVIRFLNDEHNRALLRGI
ASGRNPFGEAYGRAGDVIARKCGVPLYRFELMQTSQDIENVRKGVTEFWQRPQNA"
4858. .6414
/gene="NRDE"
4858. .6414
/gene="NRDE"
/note="similar to [SwissProt Accession Number Q08698]"
/codon_start=1
/transl_table=11
/product="RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE 2 ALPHA
CHAIN (EC 1.17.4.1) (RIBONUCLEOTIDE REDUCTASE 2)."
/protein_id="BAA16539.1"
/db_xref="GI:1800062"
/translation="MDYHALNMLNLYDSAGRIQFDKDRQAVDAFIATHVRPNSVTF
SQQRNLWLVNDEGYDESIVNRYSRDFTVLTHTAHTSGRFOTFLGAWFYTSYTLK
TDGKYLEDFADRVMTVALTLAQDETALQLTDEMLSGRFQFATFLNCGKQQRG
ELVSCFLRIEDNMESIGRAVNSALQISKRGGVAFLLSNLRAGAPIKKIENSSGV
IPVKMELEDAFSYANQLGARQAGAVILHAHPDILRFLDTKRENADEKIRIKTSLUG
WIPDITFLAKENAOALFSFYDVERVYKGFADVAISQHYDELVADEIRIKYLLNA
RDFQRLAIOFESGYPIYMYEDTVNRPVAGRIINMSNLCSILQVNSASEYDENLD
YRTGHDISCNLSLNIATMDSDFARTVETAVRGLTAVSDMSHRSVPSIAGNAA
SHAIGIGQNLHGYLAREGIAYGSPALDFTNLYFYAITWHALRTSMLARERGETTA
GFKQSRYAQ"
6272. .6964
/note="similar to [PIR Accession Number S34271]"
/codon_start=1
/transl_table=11
/product="ribonucleoside-diphosphate reductase (EC
1.17.4.1)"
/protein_id="BAA16540.1"
/db_xref="GI:1800063"
/translation="MVRRKHWSPISSIMSPGCMHCVPRCCWHANA VKPSPGNSHAM
PSGEYFQYVQLGNWQPKTAKVGLFTRSGITLPTREMAQLRDDVMRYLYNQLQAV
PPTGISYINHTASSIHPVAKVEIRKEGTGVYYPAPPTMTNENLALYQDAEIGE

gene

CDS

gene
681. .980
/gene="HLYU"
681. .980
/gene="HLYU"
/note="similar to [SwissProt Accession Number P52695]"
/codon_start=1
/transl_table=11
/product="TRANSCRIPTIONAL ACTIVATOR HLYU."
/protein_id="BAA16534.1"
/db_xref="GI:1800057"
/translation="MTELAQLQASAEQAAALLKAMSHPKELLILCMLSGSPGTSAGEL
TRTGLSASATSOHLARMREGLDSQDQRIILYSIKNEAVNAIATLKNVYCP"
complement(1561. .1965)
/gene="hnsB"
complement(1561. .1965)
/gene="hnsB"
/note="similar to [SwissProt Accession Number P30017]"
/codon_start=1
/transl_table=11
/product="DNA-BINDING PROTEIN STPA."
/protein_id="BAA16535.1"
/db_xref="GI:1800058"
/translation="MSVMQLSNIRTLRAMAREFSIDVLEEMLEKFRVVTKRRREE
EQOQRELAERQEKISTVLELMKADGINPELLGNSSAAAPRAGKKRQPRPKYKFTDV
NGTKTWTGQGRTPKPTAALAEKSLDDFLI"
complement(3120. .3476)
/gene="ygaC"
complement(3120. .>3476)
/gene="ygaC"
/note="similar to [SwissProt Accession Number P36931]
start codon is not identified yet"

gene

CDS

gene
681. .980
/gene="HLYU"
681. .980
/gene="HLYU"
/note="similar to [SwissProt Accession Number P52695]"
/codon_start=1
/transl_table=11
/product="TRANSCRIPTIONAL ACTIVATOR HLYU."
/protein_id="BAA16534.1"
/db_xref="GI:1800057"
/translation="MTELAQLQASAEQAAALLKAMSHPKELLILCMLSGSPGTSAGEL
TRTGLSASATSOHLARMREGLDSQDQRIILYSIKNEAVNAIATLKNVYCP"
complement(1561. .1965)
/gene="hnsB"
complement(1561. .1965)
/gene="hnsB"
/note="similar to [SwissProt Accession Number P30017]"
/codon_start=1
/transl_table=11
/product="DNA-BINDING PROTEIN STPA."
/protein_id="BAA16535.1"
/db_xref="GI:1800058"
/translation="MSVMQLSNIRTLRAMAREFSIDVLEEMLEKFRVVTKRRREE
EQOQRELAERQEKISTVLELMKADGINPELLGNSSAAAPRAGKKRQPRPKYKFTDV
NGTKTWTGQGRTPKPTAALAEKSLDDFLI"
complement(3120. .3476)
/gene="ygaC"
complement(3120. .>3476)
/gene="ygaC"
/note="similar to [SwissProt Accession Number P36931]
start codon is not identified yet"

Alignment Scores:

Pred. No.:	3.25e-105	Length:	19150
Score:	1248.00	Matches:	245
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	1	Gaps:	0

US-10-073-293A-4 (1-245) x D90891 (1-19150)

Qy	1	MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys	20
Db	13090	ATGGAAGCCCTACTCCACAGCCTCTCTGGTTCGGCGACTTCATGGAAGGATGCAAA	13149
Qy	21	AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla	40
Db	13150	GACAGTTTACCGATTGTATTAGTTATATTCCGGTGGCCTTGGCTCGGTCTGAATGCG	13209
Qy	41	ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly	60
Db	13210	ACCGCTCGGATTCCTCTCTCGAAGCGTTTTTTTCTCTGCTGCTCATTTATGCGAGC	13269
Qy	61	AlaSerGlnPheValIleThrAlaMetLeuAlaGlySerSerLeuTrpIleAlaAla	80
Db	13270	GCGAGCCAGTTTCGTCATTACCGCGATGCTGCAGCCGGAGTAGTTTGTGGATTGCTGCA	13329
Qy	81	LeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArg	100
Db	13330	CTGACCGTCATGGCAATGGATGTTCCCATGTGTGTATGGCCCGTCACTCGGTAGCCGT	13389
Qy	101	IleIleGlnArgLeuGlnLysSerIlyThrAlaLeuTrpAlaPheGlyLeuThrAspGlu	120
Db	13390	ATTATTCAGCGCTCTGCAAAATCGAAAACCGCCCTGTGGCGTTTGGCTGACGGATGAG	13449
Qy	121	ValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTrpSerGluAsnTrp	140
Db	13450	GTTTTTCGGCGCAACCGCAAACTGGTACGCAATAATCCGCCGTGGAGCGAAGCTGG	13509
Qy	141	MetIleGlyIleAlaPheSerSerTrpSerTrpValPheGlyThrValIleGlyAla	160
Db	13510	ATGATCGGCATTGCCTTCAGTTCATGTCATCGTGGGTATTGGTACGGTAATAGGGCA	13569
Qy	161	PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet	180
Db	13570	TTCTCCGCGCAGCGCTTCTGCAAGTTATCCCGCGTTGAAGCTGCAATTAGTTTATG	13629
Qy	181	LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys	200
Db	13630	CTTCGGGACATCTTTATGAGTTTCCTGCTCGCCTCTTTCCAGCGCAACAATCTCTTTCG	13689
Qy	201	ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla	220
Db	13690	GTTACCGCAGGTTAGTTGGTCCCTTCGACGGGTACGCATATTTCTATTCCCGTCGCC	13749
Qy	221	IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly	240
Db	13750	ATTCTGGCAGGCATTGTCTGTGGCTGCTCACTGCGTTAATCCAGGCATTCTGGCAAGGA	13809
Qy	241	AlaProAspGluLeu	245
Db	13810	GCGCCCGATGAGCTA	13824

RESULT 6

U00096_27

WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000

U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000
U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4639675

Continuation (28 of 47) of U00096 from base 2700001 (U00096 Escherichia coli K-12 MG1655)

Alignment Scores:					
Pred. No.:	2,89e-104	Length:	110000		
Score:	1248.00	Matches:	245		
Percent Similarity:	100.0%	Conservative:	0		
Best Local Similarity:	100.0%	Mismatches:	0		
Query Match:	100.0%	Indels:	0		
DB:	1	Gaps:	0		

US-10-073-293A-4 (1-245) x U00096_27 (1-110000)

Qy	1	MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys	20
Db	107639	ATGGAAGCCCTACTCCACAGCCTGCTCTGGTTCGGCGACTTCATGGAAGGATGCAAA	107698
Qy	21	AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla	40
Db	107699	GACAGTTTACCGATTGTATTAGTTATATCCGGTGGCCTTTGGCTCGGTCTGAATGCG	107758
Qy	41	ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly	60
Db	107759	ACCGCTCGGATTCTCTCTCTCGAAAGCGTTTTTTTCTCTGCATCATTTATGACGC	107818
Qy	61	AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla	80
Db	107819	GCGAGCCAGTTTCGTCAATTACCGCATGCTGGCAGCCGGAGTAGTTTGTGGATTGCTGCA	107878
Qy	81	LeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArg	100
Db	107879	CTGACCGTCATGGCAATGGATGTTGCCCATGTGTGTATGGCCCGCTCATCTCGGTAGCCGT	107938
Qy	101	IleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu	120
Db	107939	ATTATTCAGCGCTCTGCAAAATCGAAAACCGCCCTGTGGGGGTTTGGCCTGACGGATGAG	107998
Qy	121	ValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTrpSerGluAsnTrp	140
Db	107999	GTTTTTCGCCCGCAACCGCAAAACTGGTACGCAATAATCGCCGCTGGAGCGAAGCTGG	108058

AE005174_22	2200001	2310000	
AE005174_23	2300001	2410000	
AE005174_24	2400001	2510000	
AE005174_25	2500001	2610000	
AE005174_26	2600001	2710000	
AE005174_27	2700001	2810000	
AE005174_28	2800001	2910000	
AE005174_29	2900001	3010000	
AE005174_30	3000001	3110000	
AE005174_31	3100001	3210000	
AE005174_32	3200001	3310000	
AE005174_33	3300001	3410000	
AE005174_34	3400001	3510000	
AE005174_35	3500001	3610000	
AE005174_36	3600001	3710000	
AE005174_37	3700001	3810000	
AE005174_38	3800001	3910000	
AE005174_39	3900001	4010000	
AE005174_40	4000001	4110000	
AE005174_41	4100001	4210000	
AE005174_42	4200001	4310000	
AE005174_43	4300001	4410000	
AE005174_44	4400001	4510000	
AE005174_45	4500001	4610000	
AE005174_46	4600001	4710000	
AE005174_47	4700001	4810000	
AE005174_48	4800001	4910000	
AE005174_49	4900001	5010000	
AE005174_50	5000001	5110000	
AE005174_51	5100001	5210000	
AE005174_52	5200001	5310000	
AE005174_53	5300001	5410000	
AE005174_54	5400001	5510000	
AE005174_55	5500001	5528445	

Continuation (37 of 56) of AE005174 from base 3600001 (AE005174 Escherichia coli O157:H7)

Alignment Scores:			
Pred. No.:	3.58e-104	Length:	110000
Score:	1247.00	Matches:	244
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	99.6%	Mismatches:	0
Query Match:	99.9%	Indels:	0
DB:	1	Gaps:	0

US-10-073-293a-4 (1-245) x AE005174_36 (1-110000)

Qy	1	MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys	20
Db	1065	ATGGAAGCCCTACTCCACAGCCCTCCTGGTTCCGCGACCTTCATGGAAGGATGCAA	1124
Qy	21	AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla	40
Db	1125	GACAGTTACGATTGTTATTAGTTATATCCCGTGGCCCTTGGCTTCGCTGAATGCG	1184
Qy	41	ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly	60
Db	1185	ACCGCTCTGGGATCTCTCTCTCGAAAGCGTTTTTTTCTCTGTCATCTTATGCGAGC	1244
Qy	61	AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuThrIleAlaAla	80
Db	1245	GCGAGCCAGTTCTGTCATTCGCGATCTGCGACGAGGATGTTGTTGGTTGCTGCA	1304
Qy	81	LeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArg	100
Db	1305	CTGACCGTATGGCAATGGATGTTTCGCCATGTTTGTATGGCCGTCACATGCGTAGCCGT	1364
Qy	101	IleIleGlnArgLeuGlnLysSerLysThrAlaLeuThrAlaPheGlyLeuThrAspGlu	120
Db	1365	ATTATTGAGCGTCTGCAAAAATCGAAACCCCTATGGGGCTTGGCCCTGACGGATGAG	1424
Qy	121	ValPheAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTrpSerGluAsnTrp	140
Db	1425	GTTTTTGGCCCGCCACCGAAAACCTGGTACGCAATAATCCGCTGGAGCGAGAACTGG	1484

Qy	141	MetIleGlyIleAlaPheSerSerTyrPheSerSerTyrPheGlyThrValIleGlyVala	160
Db	1485	ATGATCGGCATTGCTTCAGTTTCATGTCATCGTGGGTATTTGTACGGTATAGGGCA	1544
Qy	161	PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaLeuGlyPheMet	180
Db	1545	TTCTCCGCGACGCGCTTGTCTGCAAGGTTATCCCGCGCTTGAAGCTGCATTAGGTTTTATG	1604
Qy	181	LeuProAlaLeuPheMetSerPheLeuAlaSerPheGlnArgLysGlnSerLeuCys	200
Db	1605	CTTCCGGCACTCTTATGAGTTTCCTGCTCGCTCTTTTCAGCGCAACAATCTCTTTC	1664
Qy	201	ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla	220
Db	1665	GTTACCGCAGGCTTAGTTGCTGCGCTTGCAGCGGTAAAGCTATTTTCTATTCCCGTCGCC	1724
Qy	221	IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly	240
Db	1725	ATTCTGGCAGGCATTGTCTGTGCTGCTCACTGCGTTAATCCAGGCATTCTGCGCAAGGA	1784
Qy	241	AlaProAspGluLeu	245
Db	1785	GCGCCCGATGAGCTA	1799

Sequence split into 55 fragments LOCUS BA000007 Accession BA000007

Fragment Name	Begin	End
BA000007_00	1	110000
BA000007_01	100001	210000
BA000007_02	200001	310000
BA000007_03	300001	410000
BA000007_04	400001	510000
BA000007_05	500001	610000
BA000007_06	600001	710000
BA000007_07	700001	810000
BA000007_08	800001	910000
BA000007_09	900001	1010000
BA000007_10	1000001	1110000
BA000007_11	1100001	1210000
BA000007_12	1200001	1310000
BA000007_13	1300001	1410000
BA000007_14	1400001	1510000
BA000007_15	1500001	1610000
BA000007_16	1600001	1710000
BA000007_17	1700001	1810000
BA000007_18	1800001	1910000
BA000007_19	1900001	2010000
BA000007_20	2000001	2110000
BA000007_21	2100001	2210000
BA000007_22	2200001	2310000
BA000007_23	2300001	2410000
BA000007_24	2400001	2510000
BA000007_25	2500001	2610000
BA000007_26	2600001	2710000
BA000007_27	2700001	2810000
BA000007_28	2800001	2910000
BA000007_29	2900001	3010000
BA000007_30	3000001	3110000
BA000007_31	3100001	3210000
BA000007_32	3200001	3310000
BA000007_33	3300001	3410000
BA000007_34	3400001	3510000
BA000007_35	3500001	3610000
BA000007_36	3600001	3710000
BA000007_37	3700001	3810000
BA000007_38	3800001	3910000
BA000007_39	3900001	4010000
BA000007_40	4000001	4110000
BA000007_41	4100001	4210000
BA000007_42	4200001	4310000

RESULT 10
BA000007_35
WPCOMMENT

BA000007_43	4300001	4410000			
BA000007_44	4400001	4510000			
BA000007_45	4500001	4610000			
BA000007_46	4600001	4710000			
BA000007_47	4700001	4810000			
BA000007_48	4800001	4910000			
BA000007_49	4900001	5010000			
BA000007_50	5000001	5110000			
BA000007_51	5100001	5210000			
BA000007_52	5200001	5310000			
BA000007_53	5300001	5410000			
BA000007_54	5400001	5498450			
Continuation (36 of 55) of BA000007 from base 3500001 (BA000007 Escherichia coli O157:H7)					
Alignment Scores:					
Pred. No.:	3 58e-104	Length:	110000		
Score:	1247.00	Matches:	244		
Percent Similarity:	100.0%	Conservative:	1		
Best Local Similarity:	99.6%	Mismatches:	0		
Query Match:	99.9%	Indels:	0		
DB:	1	Gaps:	0		
 US-10-073-293A-4 (1-245) x BA000007_35 (1-110000)					
Qy	1	MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys	20		
Db	33775	ATGGAAAGCCCTACTCCACAGCCTGTCTCTTGGTTGGCGACCCTTCATGGAAGCATGCAAA	33834		
Qy	21	AspSerLeuProIleValIleSerTyrlleProValAlaPheAlaPheGlyLeuAsnAla	40		
Db	33835	GACAGTTTACCGATGTATTAGTATTATCCCGTGCCCTTTCGGTTCGGTCTCAATGCG	33894		
Qy	41	ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIlelleltyrAlaGly	60		
Db	33895	ACCCGTCTGGATCTCTCTCTCGAAGGCTTTTTTCTCGCATCATTTATGACGGC	33954		
Qy	61	AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla	80		
Db	33955	GCGAGCCAGTTCGTCAITACCGCATCTGCGACGAGGAGTAGTTTGTGGTTGCTGCA	34014		
Qy	81	LeuThrValMetAlaMetAspValArgHii sValLeuTyrglyProSerLeuArgSerArg	100		
Db	34015	CTGACCGTCATGCCAATGGATGTTTCGCCATGTTTGTATGGCCCGTCACCTGCGTAGCCGT	34074		
Qy	101	IlelleGlnArgLeuGlnLysSerLyvThrAlaLeuTrpAlaPheGlyLeuThrAspGlu	120		
Db	34075	ATTATT CAGCGTCTGCAANAATCGAAGCCGCTATGGCGGTTTGGCCCTGACGAGTAG	34134		
Qy	121	ValPheAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTrpSerGluAsnTrp	140		
Db	34135	GTTTTTCGCCGCGCACCGCAAACTGGTAGCAATAATCGCCGCTGAGCGAGAACTGG	34194		
Qy	141	MetlleGlylleAlaPheSerSerTrpSerTrpValPheGlyThrValilleGlyAla	160		
Db	34195	ATGATCGGCATTTGCTTTCAGTTTCATGTCATCGTGGGTATTTCGTGTCGGTAATAGGGCCA	34254		
Qy	161	PheSerGlySerGlyLeuLeuGlnGlyTyvProAlaValGlualAlaLeuGlyPheMet	180		
Db	34255	TTCCTCCGCGAGCGCTTGCTGCAAGGTTATCCCCCGTTGAAGCTGCATTAGGTTTTATG	34314		
Qy	181	LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys	200		
Db	34315	CTTCCGCGCACTTTTATGAGTTTCTGCTCGCCTCTTTTCAGCGCAACAATCTCTTTGC	34374		
Qy	201	ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla	220		
Db	34375	GTTTACCAGCGCTTAGTTGGTGGCTTTCGACGGCTAACCGCTATTTCCTATTCCTCGCGCC	34434		
Qy	221	IlelleuAlaGlylleValCysGlyCysLeuThrAlaLeulleGlnAlaPheTrpGlnGly	240		
Db	34435	ATTCCTGGAGGCATGTCTGTGGCTGCTCTACTCGGTAAATCCAGGCANTTCTGCAAGGA	34494		
Qy	241	AlaProAspGluLeu	245		

Qy	21	AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla	40
Db	68165	GACAGTTTACCGATTGTATTAGTTATATTCCTGGCTTTCGGTTCGAGTGC	68224
Qy	41	ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly	60
Db	68225	ACCCGTCTGGGATTCTCTCTCTCGAAGCGTTTTTCTCTCGATCAATTTATGCAGGC	68284
Qy	61	AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla	80
Db	68285	GCGAGCCAGTTCGTCAATTCACGCGATGCTGGCAGCCGGAGTAGTTTGTGGTTCGTGCA	68344
Qy	81	LeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArg	100
Db	68345	CTGACCGTCATGGCAATGGATGTTTCGCATGTTTTGTATGGCCCGTCACTCGTAGCCGT	68404
Qy	101	IleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu	120
Db	68405	ATTATTACGGCTCTGCAAAATCGAAACCCGCCCTGTGGGCGTTTGGCTGACCGATGAG	68464
Qy	121	ValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgTrpSerGluAsnTrp	140
Db	68465	GTTTTTGGCGCGCAACCGCAAACTGGTACGCAATATTCGCCCGCTGGAGCGAAACTGG	68524
Qy	141	MetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIleGlyAla	160
Db	68525	ATGATCGGCATTCGCTTCAGTTTCATGGTCAATCGTGGGTATTTGGTACGTTAATAGGGCA	68584
Qy	161	PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet	180
Db	68585	TTCTCCGGCAGCGCTTCTCGAAGTTATCCCGCGTTTGAAGCAGCATTAGTTTATG	68644
Qy	181	LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys	200
Db	68645	CTTCGGGCATCTTTATGAGTTCTCTGCTCGCTCTTTCCAGCGCAAACTCTCTTTCG	68704
Qy	201	ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla	220
Db	68705	GTTACCGCAGCGTTAGTTGGTGGCCCTTTCAGCGCGTAACGCTATTTCCTATTCGCGCC	68764
Qy	221	IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly	240
Db	68765	ATTCGGCAGCGATTGCTGTGGCTGCTCACTGCGTTAATCCAGGCATTTCTGGCAGGA	68824
Qy	241	AlaProAspGluLeu	245
Db	68825	GCGCCCGATGAGCTA	68839

RESULT 12

AE005674_27

WPCOMMENT

Sequence split into 46 fragments LOCUS AE005674 Accession AE005674

Fragment Name	Begin	End
AE005674_00	1	110000
AE005674_01	100001	210000
AE005674_02	200001	310000
AE005674_03	300001	410000
AE005674_04	400001	510000
AE005674_05	500001	610000
AE005674_06	600001	710000
AE005674_07	700001	810000
AE005674_08	800001	910000
AE005674_09	900001	1010000
AE005674_10	1000001	1110000
AE005674_11	1100001	1210000
AE005674_12	1200001	1310000
AE005674_13	1300001	1410000
AE005674_14	1400001	1510000
AE005674_15	1500001	1610000
AE005674_16	1600001	1710000
AE005674_17	1700001	1810000
AE005674_18	1800001	1910000

AE005674_19	1900001	2010000
AE005674_20	2000001	2110000
AE005674_21	2100001	2210000
AE005674_22	2200001	2310000
AE005674_23	2300001	2410000
AE005674_24	2400001	2510000
AE005674_25	2500001	2610000
AE005674_26	2600001	2710000
AE005674_27	2700001	2810000
AE005674_28	2800001	2910000
AE005674_29	2900001	3010000
AE005674_30	3000001	3110000
AE005674_31	3100001	3210000
AE005674_32	3200001	3310000
AE005674_33	3300001	3410000
AE005674_34	3400001	3510000
AE005674_35	3500001	3610000
AE005674_36	3600001	3710000
AE005674_37	3700001	3810000
AE005674_38	3800001	3910000
AE005674_39	3900001	4010000
AE005674_40	4000001	4110000
AE005674_41	4100001	4210000
AE005674_42	4200001	4310000
AE005674_43	4300001	4410000
AE005674_44	4400001	4510000
AE005674_45	4500001	4607203

Continuation (28 of 46) of AE005674 from base 2700001 (AE005674 Shigella flexneri 2a str

Alignment Scores:			
Pred. No.:	8.44e-104	Length:	110000
Score:	1243.00	Matches:	243
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	99.2%	Mismatches:	0
Query Match:	99.6%	Indels:	0
DB:	1	Gaps:	0

US-10-073-293A-4 (1-245) x AE005674_27 (1-110000)

Qy	1	MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys	20
Db	83514	ATGGAAGCCCTACTCCACAGCCTCTCTCGTTCGGGACCTTCATGGAAGGATGCAAA	83573
Qy	21	AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla	40
Db	83574	GACAGTTTACCGATTGTATTAGTTATATTCGGTGGCCTTTTTCGTTCTGATGCG	83633
Qy	41	ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly	60
Db	83634	ACCCGCTGGGATCTCTCTCTCGAAAGCGTTTTTTTCTCTGCATCATTTATGAGGC	83693
Qy	61	AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla	80
Db	83694	GCGAGCCAGTTTCGTCAATTCACCGGATGCTGGCAGCCGGAGTAGTTTGTGGTTCGTCA	83753
Qy	81	LeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArg	100
Db	83754	CTGAGCGTCATGGCAATGGATGTTTCGCCATGTTTGTATGGCCCGTCACTGCGTAGCCGT	83813
Qy	101	IleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu	120
Db	83814	ATTATTACGGCTCTGCAAAATCGAAACCCGCCCTATGGGCGTTTGGCCTGACGGATGAG	83873
Qy	121	ValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgTrpSerGluAsnTrp	140
Db	83874	GTTTTTGGCGCGCCAGCCGCAAACTGGTACGCAATATTCGCCCTGGAGCGGAGAACTGG	83933
Qy	141	MetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIleGlyAla	160
Db	83934	ATGATCGGCATTCGCTTCAGTTTCATGGTCAATCGGTGATTTGGTACCGTAAATAGGGCA	83993
Qy	161	PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet	180

```

Db      83994  TTCTCCGCGACGGCTTCTGCTGCAGAGTTATCCCGCGTTGAAGCTGCATAGGTTTATG 84053
Qy      191  LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys 200
Db      84054  CTTCCGGCACCTTTATGAGTTTCTCTGCTGCCTCTTTCACGCGCAACAATCTCTTTCG 84113
Qy      201  ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla 220
Db      84114  GTTACCGCAGCTTTAGTTGGTGGCCCTTGACGCGTAACGCTATTCTTATTCCTCCGTCGCC 84173
Qy      221  IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly 240
Db      84174  ATTCTGCGAGCATTGCTGTGGCTGCTCTCACTCGGTATCCAGGCAATCTGCGAAGGA 84233
Qy      241  AlaProAspGluLeu 245
Db      84234  GCGCCCGATGAGCTA 84248

RESULT 13
AE016987
LOCUS   Shigella flexneri 2a str. 2457T section 10 of 16 of the complete
DEFINITION
ACCESSION   AE016987
VERSION     AE016987.1
KEYWORDS    GI:30042162
SOURCE      Shigella flexneri 2a str. 2457T
ORGANISM    Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Shigella.
REFERENCE   1 (bases 1 to 290380)
AUTHORS    Wei,J., Goldberg,M.B., Burland,V., Venkatesan,M.M., Deng,W.,
            Fournier,G., Mayhew,G.F., Plunkett,G. III, Rose,D.J., Darling,A.,
            Mau,B., Perna,N.T., Payne,S.M., Runyen-Janecky,L.J., Zhou,S.,
            Schwartz,D.C. and Blattner,F.R.
            Complete Genome Sequence and Comparative Genomics of Shigella
            flexneri Serotype 2a Strain 2457T
            Infect. Immun. 71 (5), 2775-2786 (2003)
            12704152
REFERENCE   2 (bases 1 to 290380)
AUTHORS    Wei,J., Goldberg,M.B., Burland,V., Venkatesan,M.M., Deng,W.,
            Fournier,G., Mayhew,G.F., Plunkett,G. III, Rose,D.J., Darling,A.,
            Mau,B., Perna,N.T., Payne,S.M., Runyen-Janecky,L.J., Zhou,S.,
            Schwartz,D.C. and Blattner,F.R.
            Direct Submission
            Submitted (13-JUN-2002) Genetics Laboratory, University of
            Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
FEATURES             source
     source          1. .290380
                    /organism="Shigella flexneri 2a str. 2457T"
                    /mol_type="genomic DNA"
                    /strain="2457T"
                    /serotype="2a"
                    /db_xref="taxon:198215"
                    complement(82..1209)
                    /gene="gcpB"
                    /locus_tag="S2733"
                    complement(82..1209)
                    /gene="gcpE"
                    /locus_tag="S2733"
                    /notes="residues 1 to 372 of 375 are 96.23 pct identical to
                    residues 1 to 372 of 372 from Escherichia coli K-12 :
                    B2515"
                    /codon_start=1
                    /transl_table=11
                    /product="hypothetical protein"
                    /protein_id="AAPI7888.1"
                    /db_xref="GI:30042163"
                    /translation="MNAQPIQRKSTRIYGVNVPIDGAPIAVQSNWTRTTDVEAT
                    VNQIKALERGVDIRVSVFTMDAAEFKLIKORVNVPLVADIFHDYRLAKVAEYGV
                    DCIRIPNGINRGNEDIRVSVFTMDAAEFKLIKORVNVPLVADIFHDYRLAKVAEYGV
                    AMRHVDLNFDPQKVSFKASDVFLAVESYRLAKIQIDPLHLGITAGGARSQAV
                    KSAIGLGLLSEGIQDGLRVSQAADPVEEIKVGFDLKLSIRSRGINFIACPTCSRQ"

     gene            /product="nucleoside diphosphate kinase"
                    /protein_id="AAPI7891.1"
                    /db_xref="GI:30042166"
                    /translation="MAIERTFSIIKPNVAKNVIGNIFARPEAAGFKIVGTXMLHLTV
                    EQARGFYAHDGKPFDFGLVEFTSGTGVIVVLEGENAVQRHDLGATNPANALACT
                    LRADYADSLITENGTHGSDSVESNAIREIYFFGEGEVCPRTR"
                    complement(4418..6517)
                    /gene="ppbC"
                    /locus_tag="S2737"
                    /notes="truncated"
                    /pseudo
                    6514..7961
                    /insertion_element"
                    /insertion_seq="ISSf14"
                    complement(6534..7730)
                    /locus_tag="S2738"

```

```

EFDVIGTVNALEQRLEDIITPMDVSIIGCVVNGPEALVSTLGTGTGNKKSGLYEDGV
RKORLDNMDIDOLEARIRAKASQOLDEARRIDVQVGKIIT"
/genes="yfgA"
/locus_tag="S2734"
complement(1236..2249)
/genes="yfgA"
/locus_tag="S2734"
complement(1236..2249)
/genes="yfgA"
/locus_tag="S2734"
/notes="putative membrane; Not classified"
/notes="residues 1 to 337 of 337 are 80.71 pct identical to
residues 1 to 337 of 337 from Escherichia coli K-12 :
B2516"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="AAPI7889.1"
/db_xref="GI:30042164"
/translation="MNTTEATHDQNEALTGTGARLNAREQLGSLQQAVERICLVKSTV
RDIEEDKAPADLASTFTLVGIRSYARLVHPIPEELFLGLEKQAPRAAKVAPMQSFSL
GKRKKRDGMLMTFTLVGLGAMWQDHKAQQEETTMADQSSAELSSNSEQ
GQSVPLNTSTTTPATTPASVDTTATNTPTAPVAPAPVDPQNAVVPQNAV
DTRATVPVTAATTDPGAAPLTPDQAGVTTTAAADPNALVMFTADCMLEVDATCKLF
SGMRKDGNNLITGQAPYKUKIGAPAAVQIQYQGVLDLSRFIRTNQVARUTLNAEQS
PAQ"
complement(2534..3688)
/genes="yfgB"
/locus_tag="S2735"
complement(2534..3688)
/genes="yfgB"
/locus_tag="S2735"
/notes="residues 1 to 384 of 384 are 99.47 pct identical to
residues 1 to 384 of 384 from Escherichia coli K-12 :
B2517"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAPI7890.1"
/db_xref="GI:30042165"
/translation="MSEQLVTPENVTTKGKINLLDNRQMBEFFKDLCEKTPRADO
VKWVHYCCDNFEDMTDINKVLRGKKEVAEIRAPEVVEEQSSDGTIKWAIAGDQ
RVETVIPEDDRATLCVSSQVGCALCEKFCSTAGQFNRLRVSIEIIGVWRAAKIVG
AAKVTQQRITVYVMVGMGEPLNANNVPAMEIMLDDFGFLSKRRVTLSTGVPFA
LDLGDMDIVALAISLHAENDEIRDEIVPINKKYNIEFLDAPVRRYLEKSNANQGRVT
IEVMDLHVNDGTGTEHAHQAELKDKPCKINLIPNPFPPDAPYGRSSNSIDRFKSVL
MSYGTIVRKTEGDDIDACGQLAGDVDRTRKTLRKMQGEALDIKAV"
complement(3838..4269)
/genes="ndk"
/locus_tag="S2736"
complement(3838..4269)
/genes="ndk"
/locus_tag="S2736"
/function="enzyme; Purine ribonucleotide biosynthesis"
/notes="residues 1 to 143 of 143 are 100.00 pct identical
to residues 1 to 143 of 143 from Escherichia coli K-12 :
B2518"
/codon_start=1
/transl_table=11
/product="nucleoside diphosphate kinase"
/protein_id="AAPI7891.1"
/db_xref="GI:30042166"
/translation="MAIERTFSIIKPNVAKNVIGNIFARPEAAGFKIVGTXMLHLTV
EQARGFYAHDGKPFDFGLVEFTSGTGVIVVLEGENAVQRHDLGATNPANALACT
LRADYADSLITENGTHGSDSVESNAIREIYFFGEGEVCPRTR"
complement(4418..6517)
/genes="ppbC"
/locus_tag="S2737"
/notes="truncated"
/pseudo
6514..7961
/insertion_element"
/insertion_seq="ISSf14"
complement(6534..7730)
/locus_tag="S2738"

```

CDS	complement (6534..7730) /locus_tag="S2738" /function="IS, phage, Tn; Transposon-related functions" /note="residues 1 to 398 of 398 are 90.95 pct identical to residues 1 to 398 of 398 from GenPept : >gb AAK18438.1 AF348706.127 (AF348706) putative transposase for IS110 [Shigella flexneri]" /codon_start=1 /transl_table=1 /product="ISSf14 orf" /protein_id="AAP17892.1" /db_xref="GI:30042167" /translation="WTRSSDVESQVFIQVGVGDTHAVAINSGKRLFDKALPNDE NKLRLSDLKQHQGILLVQDPATIGALPVAVARSEGLVGLYPLGLMRRRIADLHAG EAKTVDRAAIIAAGARTPLPHALRTLKLADEIAELSMGLGDFDDDLAAQTQASNRIR GLLTQIHPAPRVPRLPELHAPVLDLLQRYSPSEKLASLGKLLAAQCKLAPRLGKR LAADTAQALAEQTVVFTNAAAVLPLRALQLITLRKORDEVALEVEQYVLAHPLGKRVLTSMGPVGVTAAARLITEVACRAFASAAHLAAVAGLAPVTRRSGSSIRGHPRRGN KALKRALPLSAFALRDLPSRAYYTRKMSQGRHNQALIALARRCDVLFAAMRDGTF YTPQGS"
gene	7960..8442 /gene="ssea" /locus_tag="S2740" /note="truncated" /pseudo complement (9261..10037) /gene="sseb" /locus_tag="S2741" /gene="seeb" complement (9261..10037) /locus_tag="S2741" /function="phenotype; Not classified" /note="residues 1 to 258 of 258 are 94.96 pct identical to residues 4 to 261 of 261 from Escherichia coli K-12 : B2522" /codon_start=1 /transl_table=1 /product="enhancer of serine sensitivity" /protein_id="AAP17893.1" /db_xref="GI:30042168" /translation="MSEYTKNELDLLEKAAATEPAHRPAFFRTLLLESTVWVPGTAAQGE AVVEDALDQHWKEDQTVIPFTSLEALQAVDEQAFVMPVTRTFEMTLGETL FLNKLPGKFMREISLIIGSEGNPLSSQELLEGESLILSEVAREPAOMIDSLTT LFKTIKPVRAFISIKENEEAQNPLIGIADGDIERIIQATGSVATDTLPGBEPID ICKVKGKSGKISFITEHIAPIFYERRWGGFLRDFKQNRIL" 10137..10170 /note="REP (repetitive extragenic palindromic) element, contains 1 REP sequence" complement (10179..11462) /gene="pepb" /locus_tag="S2742" complement (10179..11462) /gene="pepb" /locus_tag="S2742" /function="putative enzyme; Not classified" /note="residues 1 to 427 of 427 are 92.50 pct identical to residues 30 to 456 of 456 from Escherichia coli K-12 : B2523" /codon_start=1 /transl_table=1 /product="putative peptidase" /protein_id="AAP17894.1" /db_xref="GI:30042169" /translation="MTEAMKITLSTQPADARWGERKATYSINNDGTLHLNGADDLGLI ORAAKIDGLGKHVQLSGEGWDADRCWAFQYKPGKTRKVPWPDLDDAQROQLDN RLMTIDWRDGTINAPEELGSQLAQRAVDLI SNVAGDRVTVRITKGDLDREQYMGVL HTVGRSERSPVLLALDYNPTGDKAEPYACLVGKGITFDSGYSIKOTAFWDSKMSD MGAATVGTALAFAMRGLNKKRVKFLCCADNLIISGNAPFKLGDIIITYRKNKVEWMT DAGRLVLADGLIDSAQKPEMIIDAATLTGAAKTALGNDVHALPSFDDALGRLLAS ASQENEPWTFPLAEFRHSQLPSEALNTTGSAAYPAGASTAAGFLSHFVENYQQGW LHIQCSATYRKAPVQWAGATGLGVRTINLLTA" complement (11521..11721) /gene="yfhj"
repeat_region	
gene	
CDS	
gene	
CDS	
gene	

/locus_tag="S2743"

Alignment Scores:
Pred. No.: 2.84e-103 Length: 290380
Score: 1243.00 Matches: 243
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 99.2% Mismatches: 0
Query Match: 99.6% Indels: 0
Gaps: 0
DB: 1

US-10-073-293A-4 (1-245) x AE016987 (1-290380)

Qy	1	MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys	20
Db	154513	ATGGAAGCCCTACTCACAGCCTGCTCTGTGTCGGGACCTTCATGGAAGGATGCAAA	154572
Qy	21	AspSerLeuProIleValIleSerTyrlleProValAlaPheAlaPheGlyLeuAsnAla	40
Db	154573	GACAGTTTACCGATTGTTATTAGTTATATTCGGGTGGCCTTTCGGTCTCGAATCGC	154632
Qy	41	ThrArgLeuGlyPheSerProLeuSerValPhePheSerCysIlelleTyAlaGly	60
Db	154633	ACCCGTCTGGGATTCTCTCCTCTCGAAAGCGTTTTTCTCTCGATCATTTATGACAGC	154692
Qy	61	AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla	80
Db	154693	GCGAGCCAGTTCGTTCATTACCGCGATGCTGCAGCCGGGAGTAGTTGTGGTGTCTGCA	154752
Qy	81	LeuThrValMetAlaMetAspValArgHisValLeuTyrlleGlyProSerLeuArgSerArg	100
Db	154753	CTGAGCGTCATGGCAATGGATGTTCCGCATGTTTGTATGGCCCGTCACTCGTAGCCGT	154812
Qy	101	IlelleGlnArgLeuGlnLysSerIysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu	120
Db	154813	ATTATTACGCGCTCTGCAAAAATCGAAAACCCGCTATGGGCGTTTGGCCTGACGGATGAG	154872
Qy	121	ValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgTrpSerGluAsnTrp	140
Db	154873	GTTTTTGGCGCGCGCACCGCAAAACTGGTAGCAATAATCCGCGCTGGAGCGAGAACTGG	154932
Qy	141	MetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIleGlyAla	160
Db	154933	ATGATCGGCAATTCGCTTCAGTTTCATGTCATCGTGGGTATTTTGGTACGGTAAATAGGGCA	154992
Qy	161	PheSerGlySerGlyLeuLeuGlnGlyTyrrProAlaValGluAlaAlaLeuGlyPheMet	180
Db	154993	TTCTCCGCGACGCGCTTCTGCAAGGTTATCCCGCGCTTGAAGCTGCATTAGGTTTTATG	155052
Qy	181	LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys	200
Db	155053	CTTCCGGCACCTCTTTATGAGTTTCTGCTCGCCTCTTTCAGCGCAACAACTCTTTTGC	155112
Qy	201	ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla	220
Db	155113	GTTACCGCAGCTTTAGTTGGTGCCCTTCGACGGCGTACCGCTATTCTTATTCCTCCGTCGCC	155172
Qy	221	IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly	240
Db	155173	ATTCTGGCAGGCATTGTCTGTGGCTGCTCACTGCGTTAATTCAGGCGATTCTGGCAAGGA	155232
Qy	241	AlaProAspGluLeu	245
Db	155233	CGGCCCCGATGAGCTA	155247

RESULT 14
AR386287
LOCUS AR386287 756 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 3016 from patent US 6610836.
ACCESSION AR386287
VERSION AR386287.1 GI:40096021
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

[illegible]

```

to residues 7 to 239 of 240 from E. coli K12 : B2583"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAM84500.1"
/db_xref="GI:21957654"
/translacion="MTDTFTGHTLPFAITENAVLRLQORLAANATRAVARGCRAI
RCQSLAEFRCLGDTTLPQOANSRFGCLIMFDEPLKPSNTGRLIADIILPDTHAFLWA
RTEYDPELLAAISLPDLPQVVPFVEAAPEROVINQOLPISDKPPLILLDGDGTWAEAK
KMFKPSYLLAGLPLLSLQVSHLSYQALREARQPEQHCTVEVATALLHQAGDIQAAGEL
RDHFHYRQQLAGKPHHPVGRVTTASHEENA"
complement(3721..4158)
/gene="trxC"
/note="synonym: y0919"
complement(3721..4158)
/gene="trxC"
/function="putative"
/note="residues 1 to 139 of 145 are 77.69 pct identical to
residues 1 to 139 of 139 from E. coli K12 : B2582"
/codon_start=1
/transl_table=1
/product="putative thioredoxin-like protein"
/protein_id="AAM84501.1"
/db_xref="GI:21957655"
/translacion="MNTVTCTACMATNRLPEERIDDGAKCGRGHSLFDGEVINATAET
LDKLQDLPVWIDFPAWPGPCRCSPAPIFAETAARAGKVRPVKVTAEAPALSTRF
RIRSIGITMLYGRKIDMLNGAVPKAPFDNLDELQLSRDPNS"
4395..5576
/gene="y0920"
4395..5576
/gene="y0920"
/note="residues 6 to 391 of 393 are 60.05 pct identical to
residues 1 to 345 of 345 from E. coli K12 : B2581;
residues 2 to 389 of 393 are 65.72 pct identical to
residues 1 to 364 of 366 from GenPept : >emb|CAC44352.1|
(AJ410308) YfiF protein [Pectobacterium chrysanthemi]"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAM84502.1"
/db_xref="GI:21957656"
/translacion="MWKTKWNDSFGKRGKVKVYVRSESSDDRSSNRDNRRNNSRR
DNSNRDNNSRGRSPAGNRPGDNAPRGDQNTQGNAPRNDNSRRDRSGRNERDRPSR
PARSEGSPPYDS PKMTVSRAPESEPDFDHGISGKSHIDPAQLRRORAEATR YVGNA
CKALPESRPDAI PAWYVQSVTPFRFALKMAANRKAHVVEDELA KASGPEHHGG
VCFLIKRQGLDASTYIQQQTQAKDCVLAEEVGNPNHVGIMRTCAHFGINGVLQGD
PAMLSGAAVTRTAGGAEHITAINADDFVLDITFRKAGYITVTTSSHKGVSLSUKLAL
PAKRVVLVGQESDLGTSAAWQQQGLDSVIGTGRVESVNSVATGILLAEFWMRQSAQ
G"
complement(5699..7234)
/gene="emrB"
/note="synonym: y0921"
complement(5699..7234)
/gene="emrB"
/function="putative transport; drug/analog sensitivity"
/note="multidrug resistance; residues 3 to 511 of 511 are
84.08 pct identical to residues 4 to 512 of 512 from E.
coli K12 : B2686"
/codon_start=1
/transl_table=1
/product="putative membrane translocase"
/protein_id="AAM84503.1"
/db_xref="GI:21957657"
/translacion="MAQKPLEGAPLAWMTVALSLATPMQVLDSTIANVAIPTIADGLD
SSNSQGVWTSFGVANAIGIPVTGMLAKRVGRLFLWSTGLFVLASMLCGMSNLG
MLIFRVTQGLVAGFLPGLSGLILNNYPAPKRSMLALMSMTIVVAPIFGPTLGGYI
SDNYHWGMIPIPIPIGLVVLGASTLKGRKTEIRPTDITGLVLVVGIGALCIM
LDQKELDWENSTELLVLTVAVAVITFLVWELTDHPVIDLSLPSKSNPTIGCLCL
SLQWLTFGALVLDQPLQVEGYTATWAGLASAPGVLFPVLSPLSGIRFAHRIDMKRP
LDTFSFIWYACVFWYRATPEPMDFGASAWPQFQGIACFPMTPLTITLISGLPFE
RMAAASLSMCPRTLAGISGTTITLQWTRSESHLAQLTFVFNYPNPATQRYOBL
KLGMSQQQAGAYLAREITNOGLIISANEIFWLGAQVFLMLLALIFARPDPPTSGGGG
GAH"

```

Qy 19 CysLysAspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeu 38
 Db 10561 ATTACGGATAGCTGCCCATTTGTTATCGCTGCTATTACCGTCGCTTTGGCCTC 10502
 Qy 39 AsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyr 58
 Db 10501 AGTGGGTAAACCTGGGTTTACCCCGCTGGAAGCATTTTCTTCTGCTATTATTAT 10442
 Qy 59 AlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTriPile 78
 Db 10441 GCTGGTGCACGCCAATTGTGATTACCGCCCTGCTGAGCGCGGGATGTCATTGTGGGTA 10382
 Qy 79 AlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArg 98
 Db 10381 TCGGCACCTGACGGTATGCGCAATGATGCGCCATATTTGTATGCGCCGACATTAAA 10322
 Qy 99 SerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuThrAlaPheGlyLeuThr 118
 Db 10321 CACCGCATTTTGGCAAACTATACAGGTAAACAAACGGCCTTTATGGGCTTTTGGTTAAAC 10262
 Qy 119 AspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTyrSerGlu 138
 Db 10261 GATGAGTGTTCGGCTGCAACCACTAAATTTGATGAGGATCAACGGCGCTGGAGCGAG 10202
 Qy 139 AsnTrpMetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIle 158
 Db 10201 AACTGGATGCTCGGTATCGCAGTCACCTCGTGTGTTATCTTGGGTCTTGGGAACGCCATC 10142
 Qy 159 GlyAlaPheSerGlySerGlyLeuGlnGlyTyrProAlaValGluAlaAlaLeuGly 178
 Db 10141 GCGCGGATGTCGGTAATGGGCCCTTGGAGAACTACCCCGCATTTGAAGCGTCCCTCTCC 10082
 Qy 179 PheMetLeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSer 198
 Db 10081 TTTATGCTGCTGCATTATTCTCAGTTTTTTGTTAGCTGTTTAAACGCCAATACAGC 10022
 Qy 199 LeuCysValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIlePro 218
 Db 10021 CTGACGGTGATCGCTTCGCTTCAGGTGCCTTATTAGGGGTTTATTGTTCTCAATTCCA 9962
 Qy 219 ValAlaIleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGln 235
 Db 9961 GTGECTATTTGGCCCGGATTTGCGGGTGGTGTCTCGGACTTTGCTTCAG 9911

Search completed: February 16, 2006, 19:45:57
 Job time : 6383.4 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:00:29 ; Search time 737.065 Seconds
(without alignments)
2215.341 Million cell updates/sec

Title: US-10-073-293A-4

Perfect score: 1248

Sequence: 1 MESFTQPAPGSAFMFGCK.....VCGCLTALIOAFWQAGAPDEL 245

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10073293/runat_15022006.120733.951/app_query.fasta.1
-DB=Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs07
-USER=US10073293 @CGN 1.1 900 @runat_15022006.120733.951 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -SPELOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 21.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1248	100.0	738	6	ABV75656 E. coli L
2	1248	100.0	738	14	ADZ39950 E.coli yg
3	1248	100.0	738	14	Adz36194 Escherich
4	1013.5	81.2	756	11	ACH97221 Klebsiell

5	841.5	67.4	759	10	ACF69125
6	841.5	67.4	110000	10	ACF67367_20
7	841.5	67.4	243072	10	ACF65382
8	592	47.4	801	9	ADA31112
9	426	34.1	1589	4	AAL35397
10	426	34.1	1589	4	AAL62810
11	426	34.1	1589	4	ADK43465
12	426	34.1	1589	5	ADM19434
13	426	34.1	1589	8	ABX58385
14	426	34.1	1589	12	ADJ53852
15	426	34.1	1589	12	ADJ28112
16	426	34.1	1595	4	AAL35674
17	426	34.1	1595	4	ABK43782
18	426	34.1	1595	5	ADM19672
19	426	34.1	1595	8	ABX58662
20	426	34.1	1595	12	ADJ54169
21	426	34.1	1595	12	ADJ28389
22	290.5	23.3	4947	2	AAQ86927
23	246.5	19.8	744	11	ACH95512
24	238.5	19.1	12445	2	AAI3179
25	238.5	19.1	12445	6	ABX98974
26	229.5	18.4	693	4	AAH52437
27	229.5	18.4	3058	4	AAH54221
28	229.5	18.4	3077	4	AAH54230
29	228	18.3	110000	6	ABA03041_14
30	225.5	18.1	729	6	ABN92854
31	225.5	18.1	729	13	ADS03654
32	225	18.0	110000	6	ABQ69245_14
33	223	17.9	6157	6	ABQ67195_2
34	223	17.9	6157	6	ABQ70939
35	220.5	17.7	19760	13	ADV87714
36	220.5	17.7	19760	13	ADV78967
37	220.5	17.7	110000	6	ABN71527_19
38	220.5	17.7	110000	13	ADV81204_20
39	219.5	17.6	690	6	ABN71161
40	219.5	17.6	693	6	ABN69193
41	219.5	17.6	693	13	ADV83888
42	214.5	17.2	768	10	ADC91639
43	210.5	16.9	1038	11	ABD12772
44	210.5	16.9	1086	11	ABD12814
45	204.5	16.4	714	8	ACF74375

ALIGNMENTS

RESULT 1
ABV75656
ID ABV75656 standard; DNA; 738 BP.
XX
XX
AC ABV75656;
XX
DT 29-JAN-2003 (first entry)
XX
DE E. coli L-amino acid producing gene #1.
XX
XX L-amino acid; gene; ss; E.coli.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT CDS 1..738
FT /*tag= a
XX
PN EPI239041-A2.
XX
PD 11-SEP-2002.
XX
PF 13-FEB-2002; 2002EP-00003335.
XX
PR 13-FEB-2001; 2001RU-00103865.
PR 26-FEB-2001; 2001RU-00104998.
PR 26-FEB-2001; 2001RU-00104999.
PR 28-JUN-2001; 2001RU-00117632.

PR 28-JUN-2001; 2001RU-00117633.
 XX (AJIN) AJINOMOTO CO INC.
 XX Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Gussyatiner MM;
 XX P-PSDB; ABP57756.
 DR WPI; 2002-699856/76.
 XX Novel L-amino acid producing Escherichia bacterium, is modified to
 PT enhance L-amino acid production by enhancing the activities of protein
 PT capable of making bacterium to have enhanced resistance to L-amino acids.
 XX Claim 1; Page 21-22; 33pp; English.
 XX The invention relates to a novel L-amino acid producing bacterium of
 CC genus Escherichia, modified to enhance L-amino acid production by
 CC enhancing the activities of a protein. The novel bacterium is useful for
 CC producing L-amino acid e.g. L-Thr, L-Pro, L-Leu, L-Met and L-Arg,
 CC by cultivating the bacterium in a culture medium and collecting L-amino
 CC acid to be produced and accumulated from the culture medium. The present
 CC sequence represents a gene of the invention which causes increased L-
 CC amino acid production in E. coli
 XX SQ Sequence 738 BP; 136 A; 192 C; 198 G; 212 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8.45e-124 Length: 738
 Score: 1248.00 Matches: 245
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps:
 US-10-073-293A-4 (1-245) x ABV75656 (1-738)
 QY 1 MetGluSerProThrProGluProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
 DB 1 ATGGAAAGCCCTACTCCAGCCTCTCTGTTCCGGACCTTCATGAAGGATGCAAA 60
 QY 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
 DB 61 GACAGTTTACCGATTGTTATTAGTTATATTCGGTGGCCTTTCGGTCTGAATCG 120
 QY 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
 DB 121 ACCCGTCTGGGATTCTCTCTCGAAAGCGTTTCTCTCTGTCATCAITATTATGCAAGC 180
 QY 61 AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla 80
 DB 181 GCGAGCCAGTTCTGTCATTACCCGATGCTGGCAGCCGGAGTAGTTTGGATTGCTGCA 240
 QY 81 LeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArg 100
 DB 241 CTGACCGTCATGGCAATGGATGTTCCGCATGTGTTGTATGGCCGTCCTACTGCTAGCCGT 300
 QY 101 IleIleGlnArgLeuGlnIleSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu 120
 DB 301 ATTATTACGCGTCTGCAAAATCGAAACCCGCTCTGTGGGGTTTGGCGCTGACCGATGAG 360
 QY 121 ValPheAlaAlaThrAlaIleValLeuValArgAsnAlaGlySerLeuTrpIleAlaAla 140
 DB 361 GTTTTGGCGCGCAACCGCAAACTGGTACGCAATAATCGCCGCTGGAGCGAGAACTGG 420
 QY 141 MetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIleGlyAla 160
 DB 421 ATGATCGGATTCCTTCAGTTTCATGTCATGTCGTTATTTGGTACGTAATAGGCGCA 480
 QY 161 PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet 180
 DB 481 TTCTCCGCGAGCGGCTGTGTCAGAGTTTATCCCGCGTTGAAGCTGCATTAGTTTATG 540
 QY 181 LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys 200

Db 541 CTTCGGGCACTCTTTATGAGTTTCCTGCTGCTCTTTCCAGCGCAACATCTCTTTC 600
 QY 201 ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla 220
 Db 601 GTTACCGCAGCGTTAGTTGGTGGCTTTCGAGCGCTAAACGCTATTCTTCTATTCCGTCGCC 660
 QY 221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly 240
 Db 661 ATTCTGGCAGGCAATGTCTGTGGCTGCTACTCGCTTAATCCAGGCATTCCTGCAAGGA 720
 QY 241 AlaProAspGluLeu 245
 Db 721 GCGCCCGATGAGCTA 735
 RESULT 2
 ID ADZ39950 standard; cDNA; 738 BP.
 XX AC ADZ39950;
 XX 14-JUL-2005 (first entry)
 DE E.coli ygaZ gene for improved amino acid production in bacteria.
 XX amino acid production; fermentation; transport protein; gene; ss.
 XX Escherichia coli.
 OS Key Location/Qualifiers
 FH 1..738
 FT /*tag= a
 FT /product= "amino acid biosynthesis protein"
 FT /function= "putative transport protein"
 FT /gene= "ygaZ"
 XX EP1526179-A1.
 PD 27-APR-2005.
 XX 13-FEB-2002; 2004EP-00028876.
 XX 13-FEB-2001; 2001RU-00103865.
 PR 26-FEB-2001; 2001RU-00104998.
 PR 26-FEB-2001; 2001RU-00104999.
 PR 28-JUN-2001; 2001RU-00117632.
 PR 28-JUN-2001; 2001RU-00117633.
 PR 13-FEB-2002; 2002EP-00003335.
 XX (AJIN) AJINOMOTO CO INC.
 XX Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Gussyatiner MM;
 PI WPI; 2005-308098/32.
 DR P-PSDB; ADZ39951.
 DR GENBANK; U00096.
 XX New L-amino acid producing bacterium belonging to the genus Escherichia,
 PT useful for producing L-amino acids, e.g. L-threonine, L-valine, L-
 PT proline, L-methionine, or L-arginine.
 XX Disclosure; SEQ ID NO 3; 35pp; English.
 XX The invention relates to an L-amino acid producing bacterium belonging to
 CC the genus Escherichia, where the bacterium has been modified so that the
 CC L-amino acid production by the bacterium is enhanced by enhancing
 CC activities of proteins by transformation of the bacterium with DNA coding
 CC for protein or by alteration of an expression regulation sequence of the
 CC DNA on the chromosome of the bacterium. The L-amino acid producing
 CC bacterium is useful for producing L-amino acids by fermentation,
 CC including L-threonine, L-valine, L-proline, L-methionine, or L-arginine.
 CC The genes are useful for improving L-amino acid productivity. This
 CC sequence corresponds to the ygaZ gene encoding one of the L-amino acid

CC biosynthesis pathway proteins. The protein is a putative transport
 CC protein within the biosynthesis pathway.

SQ Sequence 738 BP; 136 A; 192 C; 198 G; 212 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.45e-124 Length: 738
 Score: 1248.00 Matches: 245
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0

US-10-073-293A-4 (1-245) x AD239950 (1-738)

Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysAla 20
 Db 1 ATGGAAGCCCTACTCCACACCTGCTCTGGTCCGCGACCTTCATGGAAAGGATGCAAA 60
 Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
 Db 61 GACAGTTTACCGATTGTTATTAGTTATATTCGGGTGGCCTTTCGGTTCGGTCTGAATCG 120
 Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
 Db 121 ACCCGTCTGGGATTCTCTCTCTCGAAAGCGCTTTTCTCTGTCATCATTTATGCGAGC 180
 Qy 61 AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla 80
 Db 181 GCGAGCCAGTTCTGTCATTACCGCGATCTGCGACCGGAGTAGTTGTGGATTGCTGCA 240
 Qy 81 LeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArg 100
 Db 241 CTGACCGTCATGGCAATGATGTTCCGCACTGTGTGATGCGCCGTCACCTGCGTAGCGT 300
 Qy 101 IleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu 120
 Db 301 ATTATTACGCGTCTGCAAAAATCGAAACCGCCCTGTGGCGGTTTGGCCTGACGATGAG 360
 Qy 121 ValPheAlaAlaThrAlaLysLeuValArgAsnAenArgTTPSerGluAsnTrp 140
 Db 361 GTTTTGGCGCGCCGACCGCAAACTGTGTAGCAATATCCGCCCTGAGCGAGAACTGG 420
 Qy 141 MetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIleGlyAla 160
 Db 421 ATGATCGCATTCGCTTCAGTTTCATGTCATCGTGGGTATTTGTACGGTAATAGGGCA 480
 Qy 161 PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet 180
 Db 481 TTCTCCGCGCAGCGCTTGTCTGCAAGGTTATCCCGCGTTGAAGCTGCATTAGGTTTATG 540
 Qy 181 LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys 200
 Db 541 CTTCGGGACATTTTATGAGTTTCTGCTCGCCCTTTTCAGCGCAACATCTCTTTC 600
 Qy 201 ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla 220
 Db 601 GTTACCGCAGCGTTAGTTGGTGGCCCTTGCAGCGCTAACGCTATTTTCTATTCCGTCGCC 660
 Qy 221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly 240
 Db 661 ATTCGCGAGGCATTGTCTGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Qy 241 AlaProAspGluLeu 245
 Db 721 GCGCCCGATGAGCTA 735

RESULT 3

AD236194

ID AD236194 standard; DNA; 738 BP.

XX

AC AD236194;

XX

DT 14-JUL-2005 (first entry)
 XX Escherichia coli b2862 encoding DNA SEQ ID NO:3.
 XX amino acid production; gene; ds.

XX Escherichia coli.
 OS
 XX Key Location/Qualifiers
 XX CDS 1..738
 FT /*tag= a
 FT /product= "b2862 protein"

XX EPI526181-A1.

XX 27-APR-2005.

XX 13-FEB-2002; 2004EP-00028877.

XX 13-FEB-2001; 2001RU-00103865.

XX 26-FEB-2001; 2001RU-00104998.

XX 26-FEB-2001; 2001RU-00104999.

XX 28-JUN-2001; 2001RU-00117632.

XX 28-JUN-2001; 2001RU-00117632.

XX 28-JUN-2001; 2001RU-00117633.

XX 13-FEB-2002; 2002EP-00003335.

XX (AJIN) AJINOMOTO CO INC.

XX Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Gueyatiner MM;
 XX P-PSDB; AD236195.

XX WPI; 2005-317157/33.

XX New modified L-amino acid producing bacterium useful to enhance the
 XX production of L-amino acid by enhancing the activities of proteins in a
 XX cell of bacterium.

XX Disclosure; SEQ ID NO 3; 35pp; English.

XX The invention relates to an L-amino acid (A) producing bacterium (I)
 XX (belonging to the genus Escherichia), which is modified to enhance the
 XX production of (A) by enhancing the activities of proteins (G) or (H) in a
 XX cell of (I). Also described is a method for producing (A) comprising
 XX cultivating the bacterium in a culture medium and collecting the produced
 XX and accumulated L-amino acid. The modified bacterium has the ability to
 XX grow on a minimal medium containing L-amino acid or its analog in a
 XX minimal concentration and ability to grow faster on a medium containing L
 XX -amino acid or its analog than the unmodified strain or the wild type
 XX strain, or the parental strain of the bacterium. The present sequence
 XX encodes the E. coli b2862 protein which has L-amino acid excretion
 XX activity.

SQ Sequence 738 BP; 136 A; 192 C; 198 G; 212 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.45e-124 Length: 738
 Score: 1248.00 Matches: 245
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0

US-10-073-293A-4 (1-245) x AD236194 (1-738)

Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysAla 20
 Db 1 ATGGAAGCCCTACTCCACACCTGCTCTGGTTCGGCGACCTTCATGGAAAGGATGCAAA 60
 Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
 Db 61 GACAGTTTACCGATTGTTATTAGTTATTTCCCGTGGCCCTTTCGGTCTCGGTCGAATGCG 120
 Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60

```
Db 121 ACCGGTCGGGATTCCTCTCCCTCGAAGAGCGTTTTTCTCTCGATCATTTATGCGAGC 180
Qy 61 AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla 80
Db 181 GCGAGCCAGTTCTGTCATACCGCGATGCTGCGAGCGGGAGTAGTTTGTGGATTGCTGCA 240
Qy 81 LeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArg 100
Db 241 CTGACCGGTCATGGCAATGGATGTTCGCCATGTGTGTATGGCCGCTACTCGTAGCCGT 300
Qy 101 IleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu 120
Db 301 ATTATTACAGCTCTGCAAAATCGAAACCGCCCTGTGGGGTGTGGCTGACCGATGAG 360
Qy 121 ValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgTTrpSerGluAsnTrp 140
Db 361 GTTTTGGCGCGCAACCGCAAACTGGTACGCAATATCGCGCTGGAGCGAGAACTGG 420
Qy 141 MetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIleGlyAla 160
Db 421 ATGATCGGCATTCCTTCAGTTTCATGGTTCATGCTGTGGTATTTGGTACGGTAAATAGGGCA 480
Qy 161 PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet 180
Db 481 TTCTCCGCGACGGCTTCTCGAAGTTATCCGCGGTGTGAAGCTGCATTAAGTTTATG 540
Qy 181 LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys 200
Db 541 CTTCGGGCACCTTTATGAGTTTCTGCTCGCCTCTTTCCAGCGCAACAATCTCTTTGC 600
Qy 201 ValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAla 220
Db 601 GTTACCGCAGCGTTAGTTGGTGGCCCTTCGAGCGGTAACGCTATTTTCTATTCCCGTCGC 660
Qy 221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly 240
Db 661 ATTCTGGCAGCGATTCTCTGGTGGCTCCTCACTGGCTTAATCCAGGCATTCTGGCAGGA 720
Qy 241 AlaProAspGluLeu 245
Db 721 GCGCCCGATGAGCTA 735
RESULT 4
ACH97221
ID ACH97221 standard; DNA; 756 BP.
XX AC ACH97221;
XX AC ACH97221;
DT 29-JUL-2004 (first entry)
DE Klebsiella pneumoniae polynucleotide seqid 3016.
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
OS Klebsiella pneumoniae.
XX US6610836-B1.
XX PD 26-AUG-2003.
XX 27-JAN-2000; 2000US-00489039.
XX 29-JAN-1999; 99US-0117747P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Breton GL, Osborne M;
XX WPI; 2003-895346/82.
DR P-PSDB; ABO63670.
XX
```

```
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX Disclosure; SEQ ID NO 3016; 932pp; English.
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention
XX
SQ Sequence 756 BP; 110 A; 243 C; 229 G; 174 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1.13e-98 Length: 756
Score: 1013.50 Matches: 200
Percent Similarity: 88.3% Conservative: 18
Best Local Similarity: 81.0% Mismatches: 26
Query Match: 81.2% Indels: 3
DB: 11 Gaps: 1
```

US-10-073-293A-4 (1-245) x ACH97221 (1-756)

```
Qy 1 MetGluSerProThrProGlnProAla-----ProGlySerAlaThrPheMetGlu 17
Db 10 ATGGAAACCCCGCGCCCTTAACTGTGCCTTACCTGAACGCGTCGCCACCGTCGGCGAA 69
Qy 18 GlyCysLysAspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGly 37
Db 70 GGCCTCAAGACAGTCTCCCATCGTCATCAGCTACCTTCGGTGGCATTGCGCTCGGA 129
Qy 38 LeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIle 57
Db 130 CTCAATGTACCCGGCTGGGCTTCACCGCGTGGAAAGCCTGTCTTCTCTCGCATATT 189
Qy 58 TyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrp 77
Db 190 TACCGCGCGCCAGCAGTTTGTGATCACCCCATGTGGCCCGCGGAGTTGCTGTGG 249
Qy 78 IleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeu 97
Db 250 GTCGCGCCCTGACGCGTGATGGCGATGTCGCCACGTCGTGTATGGCCCTTCACTG 309
Qy 98 ArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeu 117
Db 310 GCGAGCGCATTCGCTCAGCGCTGGATAAAAGAAAACCGCCCTCTGGGCGCTTGGCGCTG 369
Qy 118 ThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgTrpSer 137
Db 370 ACCGATGAGGTTTTCGCGCGCGCTACCGCCAGGCTGGTTCGCGATAACCGCGCTGGAGC 429
Qy 138 GluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrVal 157
Db 430 GAAACTGGATGCTCGGTCGCTTACCTTCGTGGGCGATCGTGGGTGCGGCGACGCTG 489
Qy 158 IleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeu 177
Db 490 GCGCGCGCTGCTCGCGCACGCGCTGCTGGTGGATTATCCGGCGGTGGAAAGCGCCCTC 549
Qy 178 GlyPheMetLeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGln 197
Db 550 GGCTTTATGCTGCGCGCGCTGTTATGAGCTTCTGTGCTGGCTCTTTTCCAGGTGAGAG 609
Qy 198 SerLeuCysValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIle 217
Db 610 TCACCTCTGGTCAAGCGCGCGCTGGCGCGCGCTGGGCGCGCATCTCTGTATTATTTCCATT 669
Qy 218 ProValAlaIleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPhe 237
Db 670 CCGCGCGCGATCTCTCGCGGAAATTGCTGTGGGGGTGCTGTACGCGCGCTGCTTTCAGGCCATG 729
```

QY 238 TrpGlnGlyAlaProAspGlu 244
 Db 730 CTCAGGGGATCGCGATGAA 750

RESULT 5
 ACF69125
 ID ACF69125 standard; DNA; 759 BP.

XX AC ACF69125;
 XX 20-NOV-2003 (first entry)

XX Photorhabdus luminescens nucleotide sequence #7592.

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough; gene; ds.

XX Photorhabdus luminescens.

OS WO200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;

XX WPT; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX Claim 2; SEQ ID NO 7592; 1205pp; French.

CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes

SQ Sequence 759 BP; 195 A; 174 C; 156 G; 234 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.98e-80 Length: 759
 Score: 841.50 Matches: 160
 Percent Similarity: 82.1% Conservative: 32
 Best Local Similarity: 68.4% Mismatches: 41
 Query Match: 67.4% Indels: 1
 DB: 10 Gaps: 1

US-10-073-293A-4 (1-245) x ACF69125 (1-759)
 QY 3 SerProThrProGlnProAlaProGly---SerAlaThrPheMetGluGlyCysIysAsp 21
 Db 10 TCTGATACATCATCCCTTAAACGAGTAAATAATCTCTTTTCTAGTAAAGGAATATATGAT 69
 QY 22 SerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAlaThr 41
 Db 70 AGTTTACCCATTTGTTATCGGTTATATCCCGTCGCTTGGCTTTGGTCTCAATGCCGTC 129
 QY 42 ArgLeuGlyPheSerProLeuGlySerValPhePheSerCysIleIleTyrAlaGlyAla 61
 Db 130 AAACCTTGCTTCAACCAATGAAGCATTCTTTTCATGCATCATCTACGCCGATGCA 189
 QY 62 SerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuThrTrpIleAlaAla 81
 Db 190 ACCAGTTCGTGCATCACAGCTTTTACTGAGTCGGGGACATCATTTATGATTTCTGCCCTA 249
 QY 82 ThrValMetAlaMetAspValArgHisValIleTyrGlyProSerLeuArgSerArgIle 101
 Db 250 ACAATTATGCAATGGATGTCGCCCATATCTTTATGGTCCATCTTTAAGGCACCGTATC 309
 QY 102 IleGlnArgLeuGlnIysSerLysThrAlaLeuThrAlaPheGlyLeuThrAspGluVal 121
 Db 310 AAAGATAAGCTAACCGAGAGAAAAACCGTTATCTGGGCTTTCGGGCTTGACAGATCAAGTT 369
 QY 122 PheAlaAlaAlaThrAlaIysLeuValArgAsnAsnArgArgTyrSerGluAsnTrpMet 141
 Db 370 TTTGCCCGCCGCGACTGCAAAACCTATTAAAAACACCGAGCTGGAGTGAATCTGGATG 429
 QY 142 IleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIleGlyAlaPhe 161
 Db 430 GTTCTATTGCATCTCTTTCTTGGCTGGCTGGGCGCAGGTACCGAGCGCGTGTCATTT 489
 QY 162 SerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMetLeu 181
 Db 490 CTGGTAACGGTATTATTGGAATCTATCCGCTATAGAAGCTGCATGATTTTCATGTTA 549
 QY 182 ProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCysVal 201
 Db 550 CCAGCATATTCTCTAGTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 609
 QY 202 ThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAlaIle 221
 Db 610 GCACCGCATACACCGAGGACCTTTTAGGATTTACATTTTCTCATTTCCAGTTGCTATT 669
 QY 222 LeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGln 235
 Db 670 CTGGCAGGTATTGTGGTGGTGTGTATCGCGGCACCTGTTACAA 711

RESULT 6

ACF67367_20
 Continuation (21 of 57) of ACF67367 from base 2000001 (Photorhabdus luminescens nucleotide sequence split into 57 fragments) LOCUS ACF67367 Accession ACF67367

WP	Fragment Name	Begin	End
WP	ACF67367_00	1	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000
WP	ACF67367_03	300001	410000
WP	ACF67367_04	400001	510000
WP	ACF67367_05	500001	610000
WP	ACF67367_06	600001	710000
WP	ACF67367_07	700001	810000
WP	ACF67367_08	800001	910000
WP	ACF67367_09	900001	1010000
WP	ACF67367_10	1000001	1110000
WP	ACF67367_11	1100001	1210000
WP	ACF67367_12	1200001	1310000
WP	ACF67367_13	1300001	1410000
WP	ACF67367_14	1400001	1510000
WP	ACF67367_15	1500001	1610000
WP	ACF67367_16	1600001	1710000

WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000
WP	ACF67367_20	2000001	2110000
WP	ACF67367_21	2100001	2210000
WP	ACF67367_22	2200001	2310000
WP	ACF67367_23	2300001	2410000
WP	ACF67367_24	2400001	2510000
WP	ACF67367_25	2500001	2610000
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5648894

Alignment Scores:

Pred. No.:	3,16e-77	Length:	110000
Score:	841.50	Matches:	160
Percent Similarity:	82.1%	Conservative:	32
Best Local Similarity:	68.4%	Mismatches:	41
Query Match:	67.4%	Indels:	1
DB:	10	Gaps:	1

US-10-073-293A-4 (1-245) x ACF67367_20 (1-110000)

Qy	3	SerProThrProGlnProAlaProGly---SerAlaThrPheMetGluGlyCysLysAsp	21
Db	22141	TCTGATACATCATCCCCCTTAACGAGTAAATAATCTCTTTTACTGAAGGAATAATAGAT	22200
Qy	22	SerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAlaThr	41
Db	22201	AGTTTACCCTGTTTATCGGTATATCCCGTGGCTTTTGGTCTCAATGCCGTC	22260
Qy	42	ArgLeuGlyPheSerProLeuGluSerValPhePheSerCysAlaIleTyrAlaGlyAla	61
Db	22261	AAACTTGGCTCAACCCCAATGAAGCCATTTCTTTTCATGCATCATCGCCGTGCA	22320
Qy	62	SerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAlaLeu	81
Db	22321	AGCCAGTTGGTCATCACAGCTTACTGAGTGGGGGACATCATTTATGATTTCTGCCCTA	22380
Qy	82	ThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArgile	101
Db	22381	ACAATTATGGCAATGGATGTCGGCATATTTCTTTATGTGCTCCATCTTTAAGCACCGTATC	22440
Qy	102	IleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGluVal	121

Db	22441	AAAGATAAGCTAACCGAGAAAAAACCGTTATCTGGGCTTTCCGCTGCACAGATGAAGTT	22500
Qy	122	PheAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTyrSerGluAsnTrpMet	141
Db	22501	TTTGCCGCGGAGCTGCATAAATCATTAATAAACCCGAGCTGGAGTGAATACTGGATG	22560
Qy	142	IleGlyIleAlaPheSerSerTrpValPheGlyThrValIleGlyAlaPhe	161
Db	22561	GTTCCTATTGCAATCTGTTCTTGCTGGCTGGGCGCAGGTACCGCAGCGGTGCAATTT	22620
Qy	162	SerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMetLeu	181
Db	22621	CTTCGTAAACGGTTATTTGGAATCTATCCGCTATAGAAGTGCATGATTTTCATGTTA	22680
Qy	182	ProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCysVal	201
Db	22681	CCAGCACTATTTCAGTTTCTTCTTGCTTCTGTAGAAAAACAATAAGTATTGTGTT	22740
Qy	202	ThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAlaIle	221
Db	22741	GCAACCGCACTAACCGGAGCACTTTTAGGATTACATTTTCTCAATTCAGTTGCTATT	22800
Qy	222	LeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGln	235
Db	22801	CTGCAGGTATTGTCGGTGGTTGTATCGCGCACTGTTACAA	22842

RESULT 7

ACF65382

ID ACF65382 standard; DNA; 243072 BP.

XX ACF65382;

DT 20-NOV-2003 (first entry)

XX Photorhabdus luminescens nucleotide sequence #35.

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;

XX detection; food; gene expression; plant; animal; microorganism; toxin;

XX antibiotic; biopesticide; virulence factor; disease model; plague;

XX whooping cough; gene; ds.

XX Photorhabdus luminescens.

XX WO200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;

PI Buchrieser C;

XX WPI; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,

XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.

PS Claim 1; SEQ ID NO 35; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded

XX proteins from Photorhabdus luminescens. The isolated sequences are

XX sources of probes and primers for detecting the genome of P. luminescens

XX and related species; to study polymorphisms; for gene analysis and for

XX detection/amplification of the genes. Antibodies (Ab) raised against the

XX polypeptides encoded by the genes are used for detection/identification

XX of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that

XX carry a gene-containing vector are used to select compounds that

XX modulate, regulate, induce or inhibit expression of the genes in plants,

CC animals or microorganisms other than *P. luminescens* and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by *P.*
 CC *luminescens*. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterial useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which *P.*
 CC *luminescens* is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated *P. luminescens* genes
 XX
 SQ Sequence 243072 BP; 70273 A; 57885 C; 49427 G; 65487 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9,59e-77 Length: 243072
 Score: 841.50 Matches: 160
 Percent Similarity: 82.1% Conservatives: 32
 Best Local Similarity: 68.4% Mismatches: 41
 Query Match: 67.4% Indels: 1
 DB: 10 Gaps: 1

US-10-073-293A-4 (1-245) x ACF65382 (1-243072)

Qy 3 SerProThrProGlnProAlaProGly---SerAlaThrPheMetGluGlyCysLysAsp 21
 Db 74883 TCTGATACATCATCCCCCTTAACGAGTAAATAATCTTCTTTACTGAAGGAATAATAGAT 74942
 Qy 22 SerLeuProIleValIleSerTyIleProValAlaPheAlaPheGlyLeuAsnAlaThr 41
 Db 74943 AGTTTACCATTGTTATCGGTTATATCCCGTCGCTTGTCTTGTCTCAATGCCGTC 75002
 Qy 42 ArgLeuGlyPheSerProLeuGlySerValPhePheSerCysIleIleTyAlaGlyVala 61
 Db 75003 AACTGTGCTTCAACCCCAATGGAGCCATTTCTTTTCATGCATCATCTAGCCGGTGCA 75062
 Qy 62 SerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuThrTpIleAlaAlaLeu 81
 Db 75063 AGCCAGTCGTCATCACAGCTTTACTGAGTCGGGGACATCATTTATGGATTCTGCCCTA 75122
 Qy 82 ThrValMetAlaMetAspValArgHisValLeuTyIleGlyProSerLeuArgSerArgIle 101
 Db 75123 ACAATTATGGCAATGGATGTCCGCCATATCTTTATGTCCATCTTTAAGGCACCGTATC 75182
 Qy 102 IleGlnArgLeuGlnLysSerLysThrAlaLeuThrAlaPheGlyLeuThrAspGluVal 121
 Db 75183 AAGATAGCTAACCGGAAACCAACCGTTATCTGGGCTTTTCGGCCTGACAGATGAAGTT 75242
 Qy 122 PheAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTrpSerGluAsnTrpMet 141
 Db 75243 TTTGCCCGCGGACTGCDAAACTCATTAACCAACCCAGCGAGCTGGAGTGAACCTGGATG 75302
 Qy 142 IleGlyIleAlaPheSerSerTrpSerTrpValPheGlyThrValIleGlyAlaPhe 161
 Db 75303 GTTGCTATTGCAATCTGTTCTTGGCTGGCTGGCGCGAGTACCGCAGCGCGTGCAATT 75362
 Qy 162 SerGlySerGlyLeuLeuGlnGlyTyIleProValAlaGluAlaLeuGlyPheMetLeu 181
 Db 75363 CTTGGTAACGGTTATTGGAAATCTTATCCCGCTATAGAACTGGCATGATTTTCATGTTA 75422
 Qy 182 ProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCysVal 201
 Db 75423 CCAGCACTATTTCTCAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 75482
 Qy 202 ThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProValAlaIle 221
 Db 75483 GCAACCGCACTAACCGGAGCACTTTTAGGGATATACATTTTCTCAATTCAGTTGCTATT 75542
 Qy 222 LeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGln 235
 Db 75543 CTGGCAGGTATGTGGTGGTGTATCGCGCACTGTTTACAA 75584

RESULT 8

ADA31112
 ID ADA31112 standard; DNA; 801 BP.
 XX
 AC ADA31112;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE DNA encoding Acinetobacter baumannii protein #2399.
 XX
 KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
 KW vaccine; plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 FN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI; 2003-576092/54.
 DR P-PSDB; ADA35238.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 PS Example; SEQ ID NO 2399; 328pp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents DNA encoding an A. baumannii
 CC protein.
 XX
 SQ Sequence 801 BP; 196 A; 147 C; 165 G; 293 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.67e-53 Length: 801
 Score: 592.00 Matches: 111
 Percent Similarity: 71.0% Conservatives: 53
 Best Local Similarity: 48.1% Mismatches: 67
 Query Match: 47.4% Indels: 0
 DB: 9 Gaps: 0

US-10-073-293A-4 (1-245) x ADA31112 (1-801)

Qy 7 GlnProAlaProGlySerAlaThrPheMetGluGlyCysLysAspSerLeuProIleVal 26
 Db 76 CAGATGCTGCTCAACACAGCAACTTTTGGCAAGGTGCCAAGATAGTCAAGCCATCGTT 135
 Qy 27 IleSerTyIleProValAlaPheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSer 46
 Db 136 CTCACCTACTTACAGATTCTTTTGGCTTTTGGCGTTTCTGCATCTCAGTTGGTTTACA 195
 Qy 47 ProLeuGluSerValPhePheSerCysIleIleTyAlaGlyAlaSerGlnPheValIle 66
 Db 196 CCATGGGAACCAATTTTCTTTCTTGTTCATGTATGAGTGCAGGTGCAACCAATTTTAGTG 255
 Qy 67 ThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaLeuThrValMetAlaMet 86


```
Db      256  GTCCGACATTTTAGCAAGTGGTTCTCTATATGGTTAAACCGCTTTAAACAGTCATTGCACTT 315
Qy      87  AspValArgHisValLeuTyGlyProSerLeuArgSerArgIleIleGlnArgLeuGln 106
Db      316  GATATACACATGACTTTATATGGGCTCTCTTTATATTTAAATTCACCAATAAATTAAT 375
Qy      107  LysSerLysThrAlaLeuTyrAlaPheGlyLeuThrAspGluValPheAlaAlaThr 126
Db      376  TTAAGAAGACAGCGTTTGGCTGGGTTTACGGACGAAGTTTTTGCAGTGGTATG 435
Qy      127  AlalysLeuValArgAsnAsnArgTyrSerGluAsnTyrMetIleGlyIleAlaPhe 146
Db      436  ATCCAACATTTTACACGCGCAGACAACATGGTCCGAGTCTTGGATGTTGGGCTTAAGTCTA 495
Qy      147  SerSerTyrSerSerTyrValPheGlyThrValIleGlyAlaPheSerGlySerGlyLeu 166
Db      496  TTTAGTCGATGTCTTGGGCAAGTGGTTCATTTATAGCGGGTTATTATGGCAGATCAAGTC 555
Qy      167  LeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMetLeuProAlaLeuPheMet 186
Db      556  GCACACCTTCAAGTTTTTACAGCACGCTTTAGATTCTTATTACCTGCTTTATTTTAA 615
Qy      187  SerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCysValThrAlaAlaLeuVal 206
Db      616  AGTTTTTTATAGCGGCATTTGAACGAAGCATAGCCTTGTAGTTGCTGTATCGTTAGTG 675
Qy      207  GlyAlaLeuAlaGlyValThrLeuPheSerIleProValAlaIleLeuAlaGlyIleVal 226
Db      676  GTGTCAGCTTTGGCATGTTATTGGATTAACTATCAGCGGCTATCTTTATTGGCATTCCT 735
Qy      227  CysGlyCysLeuThrAlaLeuIleGlnAlaPhe 237
Db      736  TCAGCATTTTAGCAGGCTATTAAATATTAT 768

RESULT 9
AAL35397
ID  AAL35397 standard; cDNA; 1589 BP.
XX
AC  AAL35397;
XX
DT  08-JAN-2002 (first entry)
XX
DE  Human musculoskeletal system related polynucleotide SEQ ID NO 739.
XX
KW  Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW  antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW  vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW  cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW  neurological disease; infection; human; secreted protein;
KW  musculoskeletal system; ss.
XX
OS  Homo sapiens.
XX
PN  WO200155367-A1.
XX
PD  02-AUG-2001.
XX
PF  17-JAN-2001; 2001WO-US001338.
XX
PR  31-JAN-2000; 2000US-0179065P.
PR  04-FEB-2000; 2000US-0180628P.
PR  24-FEB-2000; 2000US-0184664P.
PR  02-MAR-2000; 2000US-0186350P.
PR  16-MAR-2000; 2000US-0189874P.
PR  17-MAR-2000; 2000US-0190076P.
PR  18-APR-2000; 2000US-0198123P.
PR  19-MAY-2000; 2000US-0205515P.
PR  07-JUN-2000; 2000US-0209467P.
PR  28-JUN-2000; 2000US-0214886P.
PR  30-JUN-2000; 2000US-0215135P.
PR  07-JUL-2000; 2000US-0216647P.
PR  07-JUL-2000; 2000US-0216880P.
PR  11-JUL-2000; 2000US-0217487P.
PR  11-JUL-2000; 2000US-0217496P.
PR  14-JUL-2000; 2000US-0218290P.
PR  26-JUL-2000; 2000US-0220963P.
PR  26-JUL-2000; 2000US-0220964P.
PR  14-AUG-2000; 2000US-0224518P.
PR  14-AUG-2000; 2000US-0224519P.
PR  14-AUG-2000; 2000US-0225213P.
PR  14-AUG-2000; 2000US-0225214P.
PR  14-AUG-2000; 2000US-0225266P.
PR  14-AUG-2000; 2000US-0225267P.
PR  14-AUG-2000; 2000US-0225268P.
PR  14-AUG-2000; 2000US-0225270P.
PR  14-AUG-2000; 2000US-0225447P.
PR  14-AUG-2000; 2000US-0225757P.
PR  14-AUG-2000; 2000US-0225758P.
PR  14-AUG-2000; 2000US-0225759P.
PR  18-AUG-2000; 2000US-0226279P.
PR  22-AUG-2000; 2000US-0226681P.
PR  22-AUG-2000; 2000US-0226868P.
PR  22-AUG-2000; 2000US-0227182P.
PR  23-AUG-2000; 2000US-0227009P.
PR  30-AUG-2000; 2000US-0228924P.
PR  01-SEP-2000; 2000US-0229287P.
PR  01-SEP-2000; 2000US-0229343P.
PR  01-SEP-2000; 2000US-0229344P.
PR  01-SEP-2000; 2000US-0229345P.
PR  05-SEP-2000; 2000US-0229509P.
PR  05-SEP-2000; 2000US-0229513P.
PR  06-SEP-2000; 2000US-0230437P.
PR  08-SEP-2000; 2000US-0231242P.
PR  08-SEP-2000; 2000US-0231243P.
PR  08-SEP-2000; 2000US-0231244P.
PR  08-SEP-2000; 2000US-0231413P.
PR  08-SEP-2000; 2000US-0231414P.
PR  08-SEP-2000; 2000US-0232080P.
PR  12-SEP-2000; 2000US-0231968P.
PR  14-SEP-2000; 2000US-0232397P.
PR  14-SEP-2000; 2000US-0232398P.
PR  14-SEP-2000; 2000US-0232399P.
PR  14-SEP-2000; 2000US-0232400P.
PR  14-SEP-2000; 2000US-0232401P.
PR  14-SEP-2000; 2000US-0233063P.
PR  14-SEP-2000; 2000US-0233084P.
PR  21-SEP-2000; 2000US-0233085P.
PR  21-SEP-2000; 2000US-0234223P.
PR  21-SEP-2000; 2000US-0234274P.
PR  25-SEP-2000; 2000US-0234997P.
PR  25-SEP-2000; 2000US-0234998P.
PR  26-SEP-2000; 2000US-0235484P.
PR  27-SEP-2000; 2000US-0235834P.
PR  27-SEP-2000; 2000US-0235836P.
PR  29-SEP-2000; 2000US-0236327P.
PR  29-SEP-2000; 2000US-0236367P.
PR  29-SEP-2000; 2000US-0236368P.
PR  29-SEP-2000; 2000US-0236369P.
PR  29-SEP-2000; 2000US-0236370P.
PR  02-OCT-2000; 2000US-0236802P.
PR  02-OCT-2000; 2000US-0237037P.
PR  02-OCT-2000; 2000US-0237038P.
PR  02-OCT-2000; 2000US-0237039P.
PR  02-OCT-2000; 2000US-0237040P.
PR  13-OCT-2000; 2000US-0239935P.
PR  13-OCT-2000; 2000US-0239937P.
PR  20-OCT-2000; 2000US-0240960P.
PR  20-OCT-2000; 2000US-0241221P.
PR  20-OCT-2000; 2000US-0241785P.
PR  20-OCT-2000; 2000US-0241786P.
PR  20-OCT-2000; 2000US-0241787P.
PR  20-OCT-2000; 2000US-0241808P.
PR  20-OCT-2000; 2000US-0241809P.
PR  20-OCT-2000; 2000US-0241826P.
```


PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0255678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-451937/48.
 DR P-PSDB; ABB03815.
 DR

PT Isolated polypeptide for treating, preventing and/or prognosing
 PT disorders related to the musculoskeletal system including musculoskeletal
 PT cancers and also for testing and detection e.g. diagnosis.
 XX

PS Claim 1; SEQ ID NO 739; 781pp + Sequence Listing; English.

XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
 CC (ABB03087-ABB04109) associated with the musculoskeletal system useful for
 CC preventing, treating or ameliorating medical conditions e.g. by protein
 CC or gene therapy. The genes are isolated from a range of human tissues
 CC disclosed in the specification. The nucleic acids, proteins, antibodies
 CC and (ant)agonists are useful in the diagnosis, treatment and prevention
 CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,

CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
 CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 1589 BP; 307 A; 412 C; 414 G; 456 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.: 2.61e-35 Length: 1589
 Score: 426.00 Matches: 84
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 34.1% Indels: 0
 DB: 4 Gaps: 0

US-10-073-293A-4 (1-245) x AAL35397 (1-1589)

Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysIys 20
 |||||
 Db 1338 ATGGAAGCCCTACTCCACAGCCCTCCCTGGTTCGGGACCTTCATGGAAGGATGCANA 1397
 |||||
 Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
 |||||
 Db 1398 GACAGTTTACCGATTGTTATTAGTTATATCCGGTGGCCCTTCGCTCGGTCTCAATGCG 1457
 |||||
 Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
 |||||
 Db 1458 ACCGTCCTGGGATTCTCTCTCGAAGCGTTTTTTTCTCTCATCATTTATGCGAGGC 1517
 |||||
 Qy 61 AlaserGlnPheValIleThrAlaMetLeuAlaGlySerSerLeuTyrIleAlaAla 80
 |||||
 Db 1518 GCGAGCCAGTTCGTCATTACCGCATCTGCGAGCCGGAGTAGTTTGTGATTGCTGCA 1577
 |||||
 Qy 81 LeuThrValMet 84
 |||||
 Db 1578 CTGACCGTCATG 1589

RESULT 10

AAI62810
 ID AAI62810 standard; cDNA; 1589 BP.
 XX
 AC AAI62810;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human cDNA SEQ ID NO 69.

XX Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.
 XX
 OS Homo sapiens.

XX WO200155449-A1.
 PN
 XX 02-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US001346.
 PF
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0216880P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250331P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-581633/65.
XX P-PSDB; AAU87135.
XX
XX New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.
XX
XX Claim 1; SEQ ID NO 55; 837pp; English.

XX The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (I) and polypeptides (II) encoded by (I), are used to treat a medical conditions and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukemia, disorders involving neovascularisation e.g. malignancies, respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g. acute kidney failure and blood related disorders e.g. myocardial infarction. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein,

Alignment Scores:

Pred. No.:	2.61e-35	Length:	1589
Score:	426.00	Matches:	84
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	34.1%	Indels:	0
DB:	4	Gaps:	0
US-10-073-293A-4 (1-245) x ABK43465 (1-1589)			
Qy	1	MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys	20
Db	1338	ATGGAAGCCCTACTCCACAGCCTCTCTGGTCCGGACCTTCATGGAAGGATGCAAA	1397
Qy	21	AspSerLeuProIleValIleSerTyrlleProValAlaPheAlaPheGlyLeuAsnAla	40
Db	1398	GACAGTTTACCGATTGTTATTAGTTATATTCGGTGGCCTTTGGCTTCGGTCTGAATGCG	1457
Qy	41	ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleIleTyAlaGly	60
Db	1458	ACCCGTCTGGATCTCTCTCTCGAAGAGCTTTTCTCTGCATCATTTATGTCAGGC	1517
Qy	61	AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla	80
Db	1518	GGAGCCAGTTCGTCAATACCGCGATGCTGCAGCCGGAGTAGTTTGTGGATTGCTGCA	1577
Qy	81	LeuThrValMet	84
Db	1578	CTGACCGTCATG	1589
RESULT 12			
ADM19434			
ID	ADM19434	standard; cDNA; 1589 BP.	
XX	AC	ADM19434;	
XX	CC		
XX	DT	20-MAY-2004 (first entry)	
XX	DE	Novel human channel/transporter gene #231.	
XX	KW	ds; gene; immunosuppressive; antiarthritic; antirheumatic;	
XX	KW	antiproliferative; cytostatic; cardiatic; vasotrophic; cerebroprotective;	
XX	KW	nocrotropic; neuroprotective; antibacterial; virucide; fungicide;	
XX	KW	ophthalmological; gene therapy; channel/transporter protein;	
XX	KW	rheumatoid arthritis; neoplasia; cardiac arrest; cerebrovascular disorder;	
XX	KW	cerebral ischemia; angiogenesis; nervous system disorder;	
XX	KW	Alzheimer's disease; ocular disorder; corneal infection; wound healing;	
XX	KW	epithelial cell proliferation; skin aging; sunburn; transplantation;	
XX	KW	chemotaxis; food additive.	
OS	Homo sapiens.		
XX	XX		
XX	FN	WO200154472-A2.	
XX	XX		
XX	PD	02-AUG-2001.	
XX	XX		
XX	PF	17-JAN-2001; 2001WO-US001307.	
XX	XX		
XX	PR	31-JAN-2000; 2000US-0179065P.	
XX	PR	04-FEB-2000; 2000US-0180628P.	
XX	PR	24-FEB-2000; 2000US-0184664P.	
XX	PR	02-MAR-2000; 2000US-0186350P.	
XX	PR	16-MAR-2000; 2000US-0189874P.	
XX	PR	17-MAR-2000; 2000US-0190076P.	
XX	PR	18-APR-2000; 2000US-0198123P.	
XX	PR	19-MAY-2000; 2000US-0205515P.	
XX	PR	07-JUN-2000; 2000US-0209467P.	
XX	PR	28-JUN-2000; 2000US-0214886P.	
XX	PR	30-JUN-2000; 2000US-0215135P.	
XX	PR	07-JUL-2000; 2000US-0216647P.	
XX	PR	07-JUL-2000; 2000US-0216880P.	
XX	PR	11-JUL-2000; 2000US-0217487P.	
XX	PR	11-JUL-2000; 2000US-0217496P.	
XX	PR	14-JUL-2000; 2000US-0218290P.	

PR	26-JUL-2000; 2000US-0220963P.
PR	26-JUL-2000; 2000US-0220964P.
PR	14-AUG-2000; 2000US-0224518P.
PR	14-AUG-2000; 2000US-0224519P.
PR	14-AUG-2000; 2000US-0225213P.
PR	14-AUG-2000; 2000US-0225214P.
PR	14-AUG-2000; 2000US-0225266P.
PR	14-AUG-2000; 2000US-0225267P.
PR	14-AUG-2000; 2000US-0225288P.
PR	14-AUG-2000; 2000US-0225270P.
PR	14-AUG-2000; 2000US-0225447P.
PR	14-AUG-2000; 2000US-0225757P.
PR	14-AUG-2000; 2000US-0225758P.
PR	14-AUG-2000; 2000US-0225759P.
PR	18-AUG-2000; 2000US-0226279P.
PR	22-AUG-2000; 2000US-0226681P.
PR	22-AUG-2000; 2000US-0226686P.
PR	23-AUG-2000; 2000US-0227182P.
PR	30-AUG-2000; 2000US-0228924P.
PR	01-SEP-2000; 2000US-0229287P.
PR	01-SEP-2000; 2000US-0229343P.
PR	01-SEP-2000; 2000US-0229344P.
PR	01-SEP-2000; 2000US-0229345P.
PR	05-SEP-2000; 2000US-0229509P.
PR	05-SEP-2000; 2000US-0229513P.
PR	06-SEP-2000; 2000US-0230437P.
PR	06-SEP-2000; 2000US-0230438P.
PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231244P.
PR	08-SEP-2000; 2000US-0231413P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	12-SEP-2000; 2000US-0231968P.
PR	14-SEP-2000; 2000US-0232397P.
PR	14-SEP-2000; 2000US-0232398P.
PR	14-SEP-2000; 2000US-0232399P.
PR	14-SEP-2000; 2000US-0232400P.
PR	14-SEP-2000; 2000US-0232401P.
PR	14-SEP-2000; 2000US-0233063P.
PR	14-SEP-2000; 2000US-0233064P.
PR	21-SEP-2000; 2000US-0233065P.
PR	21-SEP-2000; 2000US-0234223P.
PR	21-SEP-2000; 2000US-0234274P.
PR	25-SEP-2000; 2000US-0234997P.
PR	25-SEP-2000; 2000US-0234998P.
PR	26-SEP-2000; 2000US-0235484P.
PR	27-SEP-2000; 2000US-0235834P.
PR	27-SEP-2000; 2000US-0235836P.
PR	29-SEP-2000; 2000US-0236327P.
PR	29-SEP-2000; 2000US-0236367P.
PR	29-SEP-2000; 2000US-0236368P.
PR	29-SEP-2000; 2000US-0236369P.
PR	29-SEP-2000; 2000US-0236370P.
PR	02-OCT-2000; 2000US-0236802P.
PR	02-OCT-2000; 2000US-0237037P.
PR	02-OCT-2000; 2000US-0237038P.
PR	02-OCT-2000; 2000US-0237039P.
PR	02-OCT-2000; 2000US-0237040P.
PR	13-OCT-2000; 2000US-0239935P.
PR	13-OCT-2000; 2000US-0239937P.
PR	20-OCT-2000; 2000US-0240960P.
PR	20-OCT-2000; 2000US-0241221P.
PR	20-OCT-2000; 2000US-0241785P.
PR	20-OCT-2000; 2000US-0241786P.
PR	20-OCT-2000; 2000US-0241787P.
PR	20-OCT-2000; 2000US-0241808P.
PR	20-OCT-2000; 2000US-0241809P.
PR	20-OCT-2000; 2000US-0241826P.
PR	01-NOV-2000; 2000US-0244617P.
PR	08-NOV-2000; 2000US-0246474P.

08-NOV-2000; 2000US-0246475P.
 08-NOV-2000; 2000US-0246476P.
 08-NOV-2000; 2000US-0246477P.
 08-NOV-2000; 2000US-0246478P.
 08-NOV-2000; 2000US-0246523P.
 08-NOV-2000; 2000US-0246524P.
 08-NOV-2000; 2000US-0246525P.
 08-NOV-2000; 2000US-0246526P.
 08-NOV-2000; 2000US-0246527P.
 08-NOV-2000; 2000US-0246528P.
 08-NOV-2000; 2000US-0246532P.
 08-NOV-2000; 2000US-0246609P.
 08-NOV-2000; 2000US-0246610P.
 08-NOV-2000; 2000US-0246611P.
 08-NOV-2000; 2000US-0246613P.
 17-NOV-2000; 2000US-0249207P.
 17-NOV-2000; 2000US-0249208P.
 17-NOV-2000; 2000US-0249209P.
 17-NOV-2000; 2000US-0249210P.
 17-NOV-2000; 2000US-0249211P.
 17-NOV-2000; 2000US-0249212P.
 17-NOV-2000; 2000US-0249213P.
 17-NOV-2000; 2000US-0249214P.
 17-NOV-2000; 2000US-0249215P.
 17-NOV-2000; 2000US-0249216P.
 17-NOV-2000; 2000US-0249217P.
 17-NOV-2000; 2000US-0249218P.
 17-NOV-2000; 2000US-0249219P.
 17-NOV-2000; 2000US-0249220P.
 17-NOV-2000; 2000US-0249221P.
 17-NOV-2000; 2000US-0249222P.
 17-NOV-2000; 2000US-0249223P.
 17-NOV-2000; 2000US-0249224P.
 17-NOV-2000; 2000US-0249225P.
 17-NOV-2000; 2000US-0249226P.
 17-NOV-2000; 2000US-0249227P.
 17-NOV-2000; 2000US-0249228P.
 17-NOV-2000; 2000US-0249229P.
 17-NOV-2000; 2000US-0249300P.
 01-DEC-2000; 2000US-0250160P.
 01-DEC-2000; 2000US-0250391P.
 05-DEC-2000; 2000US-0251030P.
 05-DEC-2000; 2000US-0251988P.
 05-DEC-2000; 2000US-0256719P.
 06-DEC-2000; 2000US-0251479P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 08-DEC-2000; 2000US-0251989P.
 08-DEC-2000; 2000US-0251990P.
 11-DEC-2000; 2000US-0254097P.
 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-476159/51.

P-PSDB; ADM19913.

Isolated nucleic acid molecule encoding a channel/transporter protein is used in preventing, treating or ameliorating a medical condition.

Claim 1; SEQ ID NO 241; 809pp; English.

The invention relates to an isolated nucleic acid molecule encoding a channel/transporter protein or sequences at least 95% identical to a these. The nucleic acids and proteins encoded by them are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. The antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,

CC nervous system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
 CC The polypeptides can also be used to aid wound healing and epithelial
 CC cell proliferation, to prevent skin aging due to sunburn, to maintain
 CC organs before transplantation, for supporting cell culture of primary
 CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
 CC also be used as a food additive or preservative to increase or decrease
 CC storage capabilities. This sequence corresponds to a gene of the
 CC invention.
 XX
 SQ Sequence 1589 BP; 307 A; 412 C; 414 G; 456 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,61e-35 Length: 1589
 Score: 426.00 Matches: 84
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 34.1% Indels: 0
 DB: 5 Gaps: 0

US-10-073-293A-4 (1-245) x ADM19434 (1-1589)

Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
 Db 1338 ATGGAAAGCCCTACTCCACAGCCCTGCTCCCTGGTCCGCGACCTTCATGGAGGATGCATA 1397
 Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
 Db 1398 GACAGTTTACCGATTGTTATTAGTTATATTCGGTGGCCCTTGGGTTCGGTCTGATGCG 1457
 Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
 Db 1458 ACCCGTCTGGGATTCCTCCTCTCGAAAGCGGTTTTTTTCTCTGCATCATTTATGCGAGC 1517
 Qy 61 AlaSerGlnPheValIleThrAlaMetLeuAlaGlySerSerLeuTyrIleAlaAla 80
 Db 1518 GCGAGCCAGTTCGTCATTACCGGATGCTGGCAGCGGAGTAGTTTGTGATTTGCTGCA 1577
 Qy 81 LeuThrValMet 84
 Db 1578 CTGACCGTCATG 1589

RESULT 13

ABX58385

ID ABX58385 standard; cDNA; 1589 BP.

XX ABX58385;

AC ABX58385;

XX 26-FEB-2003 (first entry)

XX cDNA encoding novel human musculoskeletal system antigen #729.

XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
 XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
 XX post-operative tissue repair; limb regeneration; neuronal growth;
 XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 XX AIDS-related complex; chondrocyte growth; bone regeneration;
 XX periodontal regeneration; tissue transport; bone graft; skin aging;
 XX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 XX cell growth; organ transplant; cell differentiation; body height; weight;
 XX hair colour; eye colour; skin; percentage of adipose tissue;
 XX pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
 XX depression; tendency for violence; pain; reproductive capability;
 XX hormone level; endocrine level; appetite; libido; memory; stress;
 XX storage capability; fat content; lipid content; protein content;
 XX carbohydrate content; vitamin content; cofactor content;
 XX nutritional component.

XX Homo sapiens.

OS US2002147140-A1.

XX

```
PD 10-OCT-2002.
XX 17-JAN-2001; 2001US-00764877.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220633P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI, 2003-128199/12.
DR P-PSDB; ABU13109.
XX
XX Isolated nucleic acid molecules encoding musculoskeletal system
PT associated polypeptides, useful for detecting disorders, e.g. cancer.
XX
XX Claim 1; SEQ ID NO 739; 321pp; English.
XX
XX The invention describes an isolated nucleic acid molecule comprising a
CC sequence encoding musculoskeletal system associated polypeptides useful
CC for detecting disorders, e.g., cancer or cancer metastases, in animals or
CC humans. The nucleic acid; stimulates re-vascularisation of ischaemic
CC
```

tissues associated with conditions such as thrombosis, arteriosclerosis, and other cardiovascular conditions; treats wounds due to injuries, burns, post-operative tissue repair, and ulcers; stimulates angiogenesis and limb regeneration; stimulates neuronal growth; can treat and prevent neuronal damage occurring in certain disorders or neurodegenerative conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-related complex; stimulates chondrocyte growth, thus they can be used to enhance bone and periodontal regeneration and aid in tissue transports or bone grafts; prevents skin aging due to sunburn by stimulating keratinocyte growth; prevents hair loss, since FGF family members activate hair-forming cells and promotes melanocyte growth; stimulates growth and differentiation of hematopoietic cells and bone marrow cells when used in combination with other cytokines; maintains organs before transplantation or for supporting cell culture of primary tissues; induces tissue of mesodermal origin to differentiate in early embryos; increases or decreases the differentiation or proliferation of embryonic stem cells, besides, haematopoietic lineage; modulates mammalian characteristics, such as, body height, weight, hair colour, eye colour, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery); modulates mammalian metabolism; changes mammal's metal state or physical state by influencing biorhythms, circadian rhythms, depression, tendency for violence, tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, or stress; increases or decreases storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components. This sequence encodes a novel human musculoskeletal system antigen. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at
ftp.seqdata.uspto.gov/sequence.html?DocID=20020147140

XX SQ Sequence 1589 BP; 307 A; 412 C; 414 G; 456 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 2.61e-35 Length: 1589
Score: 426.00 Matches: 84
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.1% Indels: 0
DB: 8 Gaps: 0

US-10-073-293A-4 (1-245) x ABX58385 (1-1589)

Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db 1338 ATGGAAGCCCTACTCCACAGCCTGCTCTGGTTCGGGACCTTCATGGAAGGATGCAAA 1397
Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
Db 1398 GACAGTTTACCGATTGTATTAGTTATATCCGTCGGCTTCGTTCTGGAATCGG 1457
Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
Db 1458 ACCGTCCTGGGATTCTCTCTCTCGAAAGCGTTTTTTTCTCCTGCATCATTTATGAGGC 1517
Qy 61 AlaSerGlnPheValIleThrAlaMetIleAlaAlaGlySerSerLeuTrpIleAlaAla 80
Db 1518 GCAGAGCCAGTTTCGTCATTACCGCATGCTGCACGGAGTAGTTTGTGGATGCTGCA 1577
Qy 81 LeuThrValMet 84
Db 1578 CTGACCGTCATG 1589

RESULT 14
ADI53852
ID ADI53852 standard; cDNA; 1589 BP.
XX
AC ADI53852;
XX
DT 22-APR-2004 (first entry)
XX
DE cDNA encoding novel human protein seq id 55.
XX

KW neuroprotective; nootropic; antiparkinsonian; anticonvulsant;
KW antidiabetic; antirheumatic; antiarthritic; dermatological;
KW antiinflammatory; immunosuppressive; antithyroid; antianemic; vasotropic;
KW anti-HIV; hepatotropic; virucide; antibacterial; fungicide;
KW antiparasitic; muscular; gynaecological; gastrointestinal; respiratory;
KW cardiovascular; antiarteriosclerotic; antiarrhythmic; cardiac;
KW nephrotropic; litholytic; cytostatic; gene therapy; neural disorder;
KW Alzheimer's disease; Parkinson's disease; Huntington's chorea;
KW amyotrophic lateral sclerosis; multiple sclerosis;
KW immune system disorder; diabetes; rheumatoid arthritis;
KW systemic lupus erythematosus; autoimmune thyroiditis; haemolytic anaemia;
KW inflammatory disorder; ischaemia-reperfusion injury;
KW inflammatory bowel disease; Crohn's disease; infectious disease;
KW HIV infection; hepatitis infection; bacterial infection;
KW fungal infection; parasitic infection; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
KW cardiovascular disorder; atherosclerosis; arrhythmia; myocarditis;
KW renal disorder; acute glomerulonephritis; pyelonephritis;
KW renal lithiasis; proliferative disorder; cancerous diseases; human; gene;
KW ss.
XX Homo sapiens.
OS
XX
XX US2004018969-A1.
XX
XX 29-JAN-2004.
XX
XX 17-JAN-2001; 2001US-00764875.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198113P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-02311242P.
PR 08-SEP-2000; 2000US-02311243P.
PR 08-SEP-2000; 2000US-02311244P.
PR 08-SEP-2000; 2000US-023113P.
PR 08-SEP-2000; 2000US-0231141P.
PR 08-SEP-2000; 2000US-0231144P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 XX
 DR WPI; 2004-122079/12.
 DR P-PSDB; ADI54450.
 XX

New polypeptides and nucleic acid molecules, useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating medical conditions e.g. neural disorders, reproductive disorders or infectious diseases.
 PT
 PT
 PT
 XX
 PS Claim 1; SEQ ID NO 55; 413pp; English.
 XX
 CC The invention describes an isolated polypeptide comprising an amino acid sequence at least 90% identical to: a polypeptide fragment, domain, epitope, or full-length protein of any one of 607 amino acid sequences (I) described in the specification; a polypeptide fragment of (I), or the encoded sequence contained in (II), having biological activity; or a variant, allelic variant, or a species homologue of (I). The polypeptides and nucleic acid molecules are useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating medical conditions such as neural disorders, e.g. Alzheimer's disease, Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis or multiple sclerosis,
 CC

Alignment Scores:
 Pred. No.: 2,61e-35 Length: 1589
 Score: 426.00 Matches: 84
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 34.1% Indels: 0
 DB: 12 Gaps: 0

US-10-073-293A-4 (1-245) x ADI53852 (1-1589)

QY 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
 DB 1338 ATGGAAGGCCCTACTCCAGCCTGCTCTGCTCGCGACCTTCATCGAAGGATGCAAA 1397
 QY 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
 DB 1398 GACAGTTTACCGATTGTTATTAGTTATATTCGGTGGCCCTTTGCGTCTCGGTCTGAATGCG 1457
 QY 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleIleTyrAlaGly 60
 DB 1458 ACCCGTCTGGGATTCCTCTCGAAGCGTTTTTTTCTCTGTCATCATTTATGCGAGGC 1517
 QY 61 AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla 80
 DB 1518 GCGAGCCAGTTTCGTCATTACCGCGATGCTGGCAGCGCGAGTAGTTTGTGGATTGCTGCA 1577

QY 81 LeuThrValMet 84
 DB 1578 CTGACCGTCATG 1589
 RESULT 15
 ID ADJ28112 standard; DNA; 1589 BP.
 AC ADJ28112;
 XX 20-MAY-2004 (first entry)
 XX Human musculoskeletal system-associated contig DNA - SEQ ID 739.
 XX musculoskeletal system; cytostatic; cancer; osteoporosis;
 KW gene therapy; vaccine; human; ds; gene.
 XX Homo sapiens.
 XX US2004009488-A1.
 XX 15-JAN-2004.
 XX 13-SEP-2002; 2002US-00242515.
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234233P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234937P.
 PR 26-SEP-2000; 2000US-0234988P.
 PR 27-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 13-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-024617P.
 PR 08-NOV-2000; 2000US-024647P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249255P.
 PR 17-NOV-2000; 2000US-0249257P.

PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0251989P.
 PR 06-DEC-2000; 2000US-025179P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764877.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 XX WPI; 2004-090458/09.
 DR P-PSDB; ADJ29135.
 DR
 XX
 PT New nucleic acid molecule, useful for preparing a medicament for
 PT preventing, treating or ameliorating a medical condition e.g., cancer of
 PT musculoskeletal tissues or osteoporosis.
 XX
 PS Claim 4; SEQ ID NO 739; 289pp; English.
 XX
 CC The invention relates to a novel isolated musculoskeletal system-
 CC associated nucleic acid molecule. The nucleic acid of the invention
 CC demonstrates cytostatic and osteopathic activities and may be useful for
 CC preparing a medicament for preventing, treating or ameliorating a medical
 CC condition such as cancer of the musculoskeletal tissues or osteoporosis,
 CC possibly via gene therapy or vaccine production. The current sequence is
 CC that of the human musculoskeletal system-associated contig DNA of the
 CC invention. The current sequence is not shown within the specification per
 CC se but is available on the USPTO web-site
 CC http://seqdata.uspto.gov/sequence.html?DocID=20040009488.
 XX
 SQ Sequence 1589 BP; 307 A; 412 C; 414 G; 456 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,61e-35 Length: 1589
 Score: 426.00 Matches: 84
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 34.1% Indels: 0
 DB: 12 Gaps: 0
 US-10-073-293A-4 (1-245) x ADJ28112 (1-1589)
 Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
 Db 1338 ATGGAAGCCCTACTCCACAGCCTGCTCTGCTTCCGCGACCTTCATGGAAGGATGCAA 1397
 Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaIlePheGlyLeuAsnAla 40
 Db 1398 GACAGTTTACCGATTGTTATTAGTTATATTCGCGTGGCCCTTCGTTCTGAATGCG 1457
 Qy 41 ThrArgLeuGlyPheSerProGluSerValPhePheSerCysIleIleTyrAlaGly 60
 Db 1458 ACCCGTCGGGATTCCTCTCTCGAAGGCTTTTCTTCCTGTCATCATTTATGAGGC 1517
 Qy 61 AlaSerGlnPheValIleThrAlaMetLeuAlaGlySerSerLeuTrpIleAlaAla 80
 Db 1518 GCGAGCCAGTTCGTCTATTACCGCATGCTGCGAGCCGGAGTAGTTTGTGATTGCTGCA 1577
 Qy 81 LeuThrValMet 84
 Db 1578 CTGACCGTCATG 1589

Search completed: February 16, 2006, 18:26:10
Job time : 769.065 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:12:59 ; Search time 4989.47 Seconds
(without alignments)
2297.407 Million cell updates/sec

Title: US-10-073-293A-4

Perfect score: 1248

Sequence: 1 MESTPQAPGSAFTMEGCK.....VCGCLTALIQAFWQGPDEL 245

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10073293/runat.15022006.120736.1008/app.query.fasta.1
-DB-EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs804
-USER=US10073293 @CGN 1 1 6731 @runat.15022006.120736.1008 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hcc.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gss1.*
10: gb_gss2.*
11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1248	100.0	770	10	CL689189
C 2	564	45.2	396	11	FR0004918
C 3	224	17.9	516	9	A2130831
C 4	127	10.2	326	9	A2578984
C 5	119	9.5	1648	2	BF240758
C 6	118	9.5	473	9	AQ936633
C 7	111.5	8.9	807	9	A2211151

8	111.5	8.9	1118	10	AG396025
9	110.5	8.9	770	3	BJ574342
C 10	110.5	8.9	799	6	CF452672
C 11	109	8.7	864	10	AG127426
C 12	108.5	8.7	1067	3	BI767055
C 13	106.5	8.5	650	9	BH665038
C 14	106.5	8.5	812	9	BZ433690
C 15	106.5	8.5	1225	2	BI089645
C 16	106	8.5	1224	6	CB329965
C 17	105.5	8.5	744	6	CF445953
C 18	105	8.4	1162	8	CX117301
C 19	104.5	8.4	1163	2	BF981367
20	104	8.3	708	3	BJ576113
21	104	8.3	742	3	BJ578475
22	104	8.3	776	3	BJ566827
23	104	8.3	776	3	BJ573291
24	103.5	8.3	645	7	CK533251
25	103.5	8.3	1946	4	CNS0A5VA
C 26	103	8.3	880	9	AQ895122
C 27	103	8.3	973	7	CJ365181
C 28	103	8.3	1085	3	BM453898
C 29	103	8.3	1747	10	CL505761
C 30	102.5	8.2	775	3	BJ573793
C 31	102.5	8.2	1008	8	DN476416
32	102	8.2	689	9	CC147248
C 33	102	8.2	882	7	CV288706
34	102	8.2	922	7	CV289826
C 35	101.5	8.1	1797	10	AG279699
C 36	101	8.1	1009	4	AY109638
C 37	101	8.1	1304	8	DR121567
38	100.5	8.1	727	3	BJ573600
C 39	100.5	8.1	733	6	CF446315
C 40	100.5	8.1	746	5	BU128410
C 41	100.5	8.1	837	6	CD786623
C 42	100.5	8.1	1519	9	BZ579860
C 43	100	8.0	898	10	AG065695
C 44	100	8.0	918	10	AG072205
C 45	100	8.0	1232	2	BG426619

ALIGNMENTS

RESULT 1

LOCUS	CL689189/c	770 bp	DNA	linear	GSS 09-JUL-2004
DEFINITION	PRI0150C_C01 - PRI0150C.BR (770) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.				
ACCESSION	CL689189				
VERSION	CL689189.1	GI:50198486			
KEYWORDS	GSS.				
SOURCE	Pristionchus pacificus				
ORGANISM	Pristionchus pacificus				
REFERENCE	Rukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.				
AUTHORS	Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.				
TITLE	AppADB: an AcedB database for the nematode satellite organism				
JOURNAL	Pristionchus pacificus				
PUBLISHED	Nucleic Acids Res. 32 (1), D421-D422 (2004)				
COMMENT	14681447				

Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.

FEATURES
Location/Qualifiers


```

Qy 91 ValLeuTyrGlyProSerLeuArgSerArgile-----lleGlnArgLeuGlnLysSer 108
Db 206 GTGCTCTATTCGGCTCGATCCCAACACATCCGTCACTTCACGCTGCTCCAGAAG--- 150
Qy 109 LysThrAlaLeuTrpAlaPheGlyLeuThrAspGluValPheAlaAlaAlaThrAlaLys 128
Db 149 -----TTCTTCGCCTTCTCTCTGCTGATCGATCGCAATATGCGGAGCGAAGG 99
Qy 129 LeuValArgAsnAsnArgArgTrpSerGluAsnTrpMetIleGlyIle-AlaPheSerSe 148
Db 98 CGGCGGAGCGGGCTTCGGCTCAGCTTTCTCTGCTATCTCGGCTTCGGCTCGGTGATC 39
Qy 148 rTrpSerSerTrpValPheGlyThrValIleGlyAla 160
Db 38 TATTTCCTCCCTGGCTCGTCAACACACGCGGACCGGCGCA 2

RESULT 5
BF240758/6 1648 bp mRNA linear EST 14-NOV-2000
LOCUS 601875351F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4091893 5',
DEFINITION mRNA sequence.
ACCESSION BF240758
VERSION BF240758.1 GI:11154682
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1648)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-romail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC948 row: e column: 14
High quality sequence stop: 484.
Location/Qualifiers
1. 1648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4091893"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 55"
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
site_1: SfiI (ggcgccgcggcc); Site_2: SfiI
(ggcattatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGCACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

ORIGIN

Alignment Scores:

```

Pred. No.: 0.0973 Length: 1648
Score: 119.00 Matches: 72
Percent Similarity: 41.1% Conservative: 36
Best Local Similarity: 27.4% Mismatches: 91
```

```

Query Match: 9.5% Indels: 64
DB: 2 Gaps: 13
US-10-073-293A-4 (1-245) x BF240758 (1-1648)
Qy 13 AlaThrPheMetGluGlyCysLysAspSerLeuProIleValIleSerTyrIleProVal 32
Db 1626 TCGGTTTTTCTTTCTGGCTGC-----TTTCGTTGGTTTACTTAATGTT 1585
Qy 33 AlaPheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeu----- 48
Db 1584 TCTGCTGATTCCTCGCTCTCTTCGACGTTTGGCAATCTTACCTATTCTCTACCGGCT 1525
Qy 49 -----GluSerValPhePheSer-CysIleIleTyrAlaGlyAlaSerGlnPhe-- 64
Db 1524 CCCTCTTCTCATGACTGCTTCTTATCGTTGCTCATTTACGCGTCTGTATATTCTC 1465
Qy 65 ----ValIleThr-AlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla---LeuT 82
Db 1464 TCTGTTCTCTCTTCTCTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCTTC 1405
Qy 82 hrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArgIleI 102
Db 1404 TCGTACTTTCGCTCAGTCTCTTCTTCGTTGTTCTAT-----CCTTGGCGCTTCGTTCC 1351
Qy 102 leGlnArgLeuGlnLysSerIlyThrAlaLeuTrpAlaPheGlyLeuThrAspGluValP 122
Db 1350 TGTCTGCGCTCTCCCTATTGAGAGCGCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCT 1300
Qy 122 heAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTrpSer---GluAsnTrpM 141
Db 1299 TCTCATCGCTCGCT---CTGCTGCTCGCTGCTCTCTCTGCTGCTGCTCTCCAGTGGT 1243
Qy 141 etile-----GlyIleAlaPheSerSerTrpSerSerTrpValPheGlyT 156
Db 1242 GCCTCTTGGACCCCTCTCTCTCTGCTGCGCTCTCTCTATGTCCTGCTCTCTCTCTCTG 1183
Qy 156 hrValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAla 176
Db 1182 TCGGCGCTTCTGCTCGGGTCTCGGCTCGCTCGCT-----CCGTTGCTTCCTATG 1129
Qy 176 laLeuGlyPheMetLeuProAlaLeuPheMetSerPhe-LeuLeuAla----- 191
Db 1128 TTCGGGCTCTC-----TCTCTCTTTGTTGTTGCTCTCTGTCGCGTCTCTTTCG 1078
Qy 192 -----SerPheGlnArgLysGlnSerLeuCysValThrAlaAlaLeuVal 206
Db 1077 CCGGTGCTCTGACGCTCTCTCTGCTGCTTTCGAGCTTCGCGCTTCGCTGCTGCTGCTC 1018
Qy 207 Gly-AlaLeuAlaGlyValThrLeu----- 214
Db 1017 GGTGTGCTAGTCTGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 958
Qy 215 -----PheSerIleProValAlaIleLeuAlaGlyIleValCysG 228
Db 957 CGCATATCTGTCGGTGTAGTTACATCGTGGTGGTCTCTTGTCTGCTGTGTGTGTG 898
Qy 228 Y 228
Db 897 A 897

RESULT 6
AQ936633 473 bp DNA linear GSS 23-AUG-2000
LOCUS HSJ41-905 Human NotI clones Homo sapiens genomic, genomic survey
DEFINITION sequence.
ACCESSION AQ936633
VERSION AQ936633.1 GI:7213011
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```



```

Qy 131 ArgAsnAsnArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSer 150
Db 498 -----:-----:-----:-----:-----:-----:-----:
Qy 151 SerTrpValPheGlyThrValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyr 170
Db 525 ---TGGGTTGTTGGTTGTCGGGG-----GGCTCGGG---CTGTCGGGTTT 569
Qy 171 ProAlaValGluAlaAlaLeuGly-----PheMetLeuProAlaLeuPheMet 186
Db 570 GCGGCGCGCCCTGCCTCTGGCGTGGCGTGGCTTCTGCTTTCCTCCGCTGTGTTTTTC 629
Qy 187 SerPhe-LeuLeu-----:-----:-----:-----:AlaSerPheGlnAr 195
Db 630 TCTTCCCTCTCTGGCTCTCTTTTGGCTGTGGTGACACTCGACCTGGCTTTTCTCGG 689
Qy 195 GlyGlnSerLeuGlyValThrAlaAlaLeu---ValGlyAlaLeuAlaGlyValThrLe 214
Db 690 GCGGCGGTGCTGGCGCTGGCGGTGGCTGGGGCGCGGTGGCGGGTGTGTGTG 749
Qy 214 uPheSerIleProValAlaIleLeuAlaGlyIleValCysGlyCysLeu 230
Db 750 CCGCGCGCTGCC-----:-----:-----:TTGGTCGCGCGGTGTG 780

RESULT 8
AG396025 1118 bp DNA linear GSS 21-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-215P11.T7, genomic survey
DEFINITION
ACCESSION AG396025
VERSION AG396025.1 GI:48028854
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and
Shiroishi,T.
Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
15574823
2 (bases 1 to 1118)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kunya Abe (abe@crc.riken.jp).
Tsukuba Institute of Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@crc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY : pBACe3.6
Vector : EcorI.
R.site 1 : EcorI.
R.site 2 : EcorI.
FEATURES
location/Qualifiers
1..1118
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
source

```

```

/db_xref="taxon:57486"
/clone="MSMg01-215P11.T7"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

```

ORIGIN

```

Alignment Scores:      Length: 1118
Pred. NO.:            Matches: 70
Score:                Conservative: 36
Percent Similarity:   38.7%
Best Local Similarity: 25.5%
Query Match:          Mismatches: 98
                        Indels: 70
                        Gaps: 13

```

US-10-073-293A-4 (1-245) x AG396025 (1-1118)

```

Qy 4 ProThrProGlnProAlaPProGlySerAla-----ThrPheMetGluGly 18
Db 55 CCTCTCCCGCGCGCCACCGCGCTTGTCTCCCGCGCGCGCTCGNTCGTTTGGG 114
Qy 19 CysLysAspSerLeuProIleValIleSerTyrIleProValAla-----PheAla 35
Db 115 GGGGTGCTCTCTCCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 174
Qy 36 PheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSer 54
Db 175 TTGGCCTCCCGGG-----CGTTGCGGTTTTTCGCTCTTTTGTGTGTGTGTGTGTG 231
Qy 55 -----CysIleIleTyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAla 71
Db 232 TCCGTAAGTCTGGGTCTCTTGGGTCTCTTGGGTCTGTGTGTGTGTGTGTGTGTGT 291
Qy 72 AlaGlySerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAsp-ValArgHisVa 91
Db 292 TTGTTGTGTTCATGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 346
Qy 91 lLeuTyrGlyProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAl 111
Db 347 -----CTCAGTGGCGGTCTT-----:-----:-----:ACTGTGCGCG 372
Qy 111 aLeuTrpAlaPheGlyLeuThrArgGluValPheAlaAlaAlaThrAlaLysLeuValAr 131
Db 373 TGGTCTCGGTTT-----:-----:-----:TTGCTGCCCTCTGGGCTGCTCTTTTTCG 414
Qy 131 gAsnAsnArgArgTrpSerGluAsnTrpMetIleGlyIleAla-PheSerSerTrp---- 149
Db 415 TTTCGTTTGTGGCTGGTTTGTGTGTGGGGGTTTGTGTGTGTGTGTGTGTGTGTGT 474
Qy 150 --SerSerTrpValPheGlyThrValIleGlyAlaPheSer-GlySerGlyLeuLeuGln 168
Db 475 TTTCGGGTGGGTGGTTTCTTTGTCTTGTCCCGCGTGGGCGCGGTGGTGGTGGTGG 531
Qy 169 GlyTyrProAlaValGluAlaLeuGlyPheMetLeuProAlaLeuPheMetSerPhe 188
Db 532 -----:-----:-----:-----:-----:-----:-----: 573
Qy 189 LeuLeuAlaSerPheGlnArgLysGlnSerLeuGlyValThrAlaAlaLeuValGlyAla 208
Db 574 -----TCTTTTCTCGC---TGTGGGCTTTGTGTGTGGGCGCTGGGTGTCCGCG 621
Qy 209 LeuAlaGlyValThrLeuPheSerIleProValAlaIleLeuAlaGlyIle----- 225
Db 622 GCTGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681
Qy 226 -----:-----:-----:-----:-----:-----:-----: 232
Db 682 CTTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 741
Qy 233 Leu-----:-----:-----:-----:-----:-----:-----: 238
Db 742 CTGTGTGCTGTGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 777

```


RESULT 9

BJ574342 770 bp mRNA linear EST 18-DEC-2002
 LOCUS BJ574342 Ipomoea nil mixture of flower and flower bud Ipomoea nil
 DEFINITION cDNA clone jn26102 3', mRNA sequence.

ACCESSION BJ574342
 VERSION BJ574342.1 GI:27256170

KEYWORDS

SOURCE EST.

ORGANISM Ipomoea nil (Japanese morning glory)

REFERENCE 1 (bases 1 to 770)
 AUTHORS Hoshino, A., Seki, M., Shin-i, T., Carninci, P., Kamiya, A., Shiraki, T., Nitasaka, E., Shinozaki, K., Hayashizaki, Y., Kohara, Y. and Iida, S.
 TITLE ESTs of Japanese morning glory
 JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES

Source 1. .770
 Location/Qualifiers
 1. .770
 /organism="Ipomoea nil"
 /mol_type="mRNA"
 /cultivar="Tokyo-kokei standard"
 /db_xref="taxon:35883"
 /clone="jn26102"
 /tissue_type="mixture of flower and flower bud"
 /clone_lib="Ipomoea nil mixture of flower and flower bud"

ORIGIN

Alignment Scores:
 Pred. No.: 0.282 Length: 770
 Score: 110.50 Matches: 52
 Percent Similarity: 36.1% Conservative: 23
 Best Local Similarity: 25.0% Mismatches: 74
 Query Match: 8.9% Indels: 59
 DB: 3 Gaps: 7

US-10-073-293A-4 (1-245) x BJ574342 (1-770)

Qy 3 SerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLysAspSer 22
 Db 166 ACCCTTAATCAAAACCGTATCCGCCACCTATCTATCTACCATCTTCCACATCACTTC 225
 Qy 23 LeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAlaThrArg 42
 Db 226 CTCCTCCCGCGTGCCTTCTTCTCCCGGC---GACTTGACGGAC 270
 Qy 43 LeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyr----- 58
 Db 271 TTGGCTTCAGCCCTTACGGCGAGCCTTTTACAGGAGTAGCTTCTTCGGCGGACCC 330
 Qy 59 -----AlaGlyAlaSerGlnPheValIleThrAlaMetLeuAla 71
 Db 331 TTGGCAGGAGGAGCGCTTGCTCTCGGGTTGATCGGGTGGCGCTCTCTGGCAACCTTAGCC 390
 Qy 72 AlaGlySerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisVal 91
 Db 391 GCGCGCTCTCTTCTTGGCTTGGTGGCGAGCCTTAGTCTTCGACGTGCTCTTA----- 441
 Qy 92 LeuTyGlyProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAla 111
 Db 442 -----GCTTAGCCG 450
 Qy 112 LeuTrpAlaPheGlyLeuThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArg 131
 Db 451 GCAGCGCCCTTTGGCTTA-----GCGGCAACAGCC----- 480

Qy 132 AsnAsnArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerSer 151
 Db 481 -----TTCCGCTTTGGAGCTGGATTAGCC 504
 Qy 152 TrpValPheGlyThrValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrPro 171
 Db 505 TTAGCAGCGGGCTTCGCTTCGGCTTGGCGCGGCTTTGGCTTGGCAGCAGCCTTG-- 561
 Qy 172 AlavalGluAlaAlaLeuGlyPheMetLeuProAlaLeuPheMetSerPheLeuLeuAla 191
 Db 562 GCCTTAGCTTAGCTTGGCTTGGCAGGGGAGCAGCCTTCTTCTCTCTGGTAGCCTTT 621
 Qy 192 SerPheGlnArgLysGlnSerLeu 199
 Db 622 GGCTTCGCCCGCGCTCGCAGGCTT 645

RESULT 10

CF452672/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Source

Location/Qualifiers

1. .799

/organism="Allium cepa"

/mol_type="mRNA"

/cultivar="Red Creole (bulbs), unknown (callus), Ebano &

Texas Legend (roots)"

/db_xref="taxon:4679"

/clone="ACAEM85"

/tissue_type="Callus, roots, and young bulbs"

/clone_lib="normalized cDNA library of onion"

/notes="Vector: pCMVSPORT6.1-cdb (Invitrogen); Site 1:

ECORV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA

from callus, roots, and young bulbs were combined to

synthesize the library. Normalization to enrich for

low-copy transcripts was performed by proprietary

techniques of Invitrogen."

ORIGIN

Alignment Scores:

Pred. No.: 0.297 Length: 799

Score: 110.50 Matches: 63

Percent Similarity: 43.6% Conservative: 29

Best Local Similarity: 29.9% Mismatches: 69

Query Match: 8.9% Indels: 51

DB: 6 Gaps: 12

US-10-073-293A-4 (1-245) x CF452672 (1-799)

genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Alignment Scores:

Pred. No.: 0.606 Length: 650
Score: 106.50 Matches: 45
Percent Similarity: 44.2% Conservative: 23
Best Local Similarity: 29.2% Mismatches: 54
Query Match: 8.5% Indels: 32
DB: 9 Gaps: 7

US-10-073-293A-4 (1-245) x BH665038 (1-650)

Qy 62 SerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAlaLeu 81
Db 538 TCACGTGTTTATCTTAGGCTGTAAACATCAGCTGGTTCCTCCGAGCTAATCCGGGTGTCA 479
Qy 82 ThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArgile 101
Db 478 TCTCTTTCACCTACGACATCTACCGAACATACATAATCTTAGAGCAACCGGAAA--- 422
Qy 102 IleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGluVal 121
Db 421 ---CAAATCTTGAAGTCTCAAGATCCGCGTCTTCTGGATTGGATGTTTCATCGGGATT 365
Qy 122 PheAlaAlaIleThrAlaLysLeuValArgAsnAsnArgTrpSerGluAsnTrpMet 141
Db 364 CTTGCA-----GTGATCTTGAACAAAGCCGGTGTTCCTCCGATGGATG 320
Qy 142 -----ileGlyIleAlaPheSer---Ser 148
Db 319 TATCTCGGATGGGAGTTCTCATTTGGCTCTCGGTTATTCCTCCATCGCTTCATGCTTTTG 260
Qy 149 TrpSer---SerTrpValPheGlyThrValIleGlyAlaPheSerGlySer-----Gly 165
Db 259 TGGAGCAAGACTAATTCATTTGGTGTATCTTGGTGCATCTCTGGCTGTGTCTTCTGGG 200
Qy 166 Leu-LeuGlnGlyTyrPro-----AlaValGluAlaAlaLe 177
Db 199 ATCGTTACTTGGTTATCCACCGCAAGAAAGTTCAGTACGACGCTGTGTGATCTTGATACCACA 140
Qy 177 uGlyPheMetLeuProAlaLeuPheMetSerPheLeuLeu 190
Db 139 GGGAGAAATGCTCCCATCTGCTGGTAACTCTGTTGCTA 100

RESULT 14

BZ433690/c
LOCUS BZ433690 812 bp DNA linear GSS 13-DEC-2002
DEFINITION BONPR54TF BO 1.6.2 KB tot Brassica oleracea genomic clone BONPR54,
genomic survey sequence.

ACCESSION BZ433690.1 GI:26683096

VERSION GSS.

KEYWORDS Brassica oleracea

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

Other GSSs: BONPR54TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF
Class: sheared ends.

FEATURES

source location/Qualifiers
1..812
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clones="BONPR54"
/clone_lib="BO 1.6.2 KB tot"
/note="Vector: pHOS1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Alignment Scores:
Pred. No.: 0.829 Length: 812
Score: 106.50 Matches: 45
Percent Similarity: 44.2% Conservative: 23
Best Local Similarity: 29.2% Mismatches: 54
Query Match: 8.5% Indels: 32
DB: 9 Gaps: 7

US-10-073-293A-4 (1-245) x BZ433690 (1-812)

Qy 62 SerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAlaLeu 81
Db 473 TCACGTGTTTATCTTAGGCTGTAAACATCAGCTGGTTCCTCCGAGCTAATCCGGGTGTCA 414
Qy 82 ThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArgile 101
Db 413 TCTCTCTTCACTACGACATCTACCGAACATACATAATCTTAGAGCAACCGGAAA--- 357
Qy 102 IleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGluVal 121
Db 356 ---CAAATCTTGAAGTCTCAAGATCCGCGTCTTCTGGATTGGATGTTTCATCGGGATT 300
Qy 122 PheAlaAlaIleThrAlaLysLeuValArgAsnAsnArgTrpSerGluAsnTrpMet 141
Db 299 CTTGCA-----GTGATCTTGAACAAAGCCGGTGTTCCTCCGATGGATG 255
Qy 142 -----ileGlyIleAlaPheSer---Ser 148
Db 254 TATCTCGGATGGGAGTTCTCATTTGGCTCTCGGTTATTCCTCCATCGCTTCATGCTTTTG 195
Qy 149 TrpSer---SerTrpValPheGlyThrValIleGlyAlaPheSerGlySer-----Gly 165
Db 194 TGGAGCAAGACTAATTCATTTGGTGTATCTTGGTGCATCTCTGGCTGTGTCTTCTGGG 135
Qy 166 Leu-LeuGlnGlyTyrPro-----AlaValGluAlaAlaLe 177
Db 134 ATCGTTACTTGGTTATCCACCGCAAGAAAGTTCAGTACGACGCTGTGTGATCTTGATACCACA 75
Qy 177 uGlyPheMetLeuProAlaLeuPheMetSerPheLeuLeu 190
Db 74 GGGAGAAATGCTCCCATCTGCTGGTAACTCTGTTGCTA 35

RESULT 15

BIO89645/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

Other GSSs: BIO89645

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

BIO89645 1225 bp mRNA linear EST 20-JUN-2001
DEFINITION 602854335F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4995824 5',
mRNA sequence.

ACCESION BIO89645

VERSION BIO89645.1 GI:14507975

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

Other GSSs: BIO89645

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:27:54 ; Search time 223.666 Seconds
(without alignments)
1947.114 Million cell updates/sec

Title: US-10-073-293A-4

Perfect score: 1248

Sequence: 1 MESPTQPAPGSAFMEGCK.....VCGLTALIQAFQWQAPDEL 245

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2506114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/abs/ABSSWEB.spool/US10073293/runat.15022006.120738.1057/app_query.fasta_1
-DB=Issued_Patents_NA -OFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs08 -USER=US10073293.0CGN 1.1 237 @runat.15022006.120738.1057 -NCPU=6
-ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/1.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*
8: /cgn2_6/ptodata/1/ina/R.COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1013.5	81.2	756	3	US-09-489-039A-3016
2	592	47.4	801	3	US-09-328-352-2399
3	246.5	19.8	744	3	US-09-489-039A-1307
4	229.5	18.4	693	3	US-09-710-279-267
5	229.5	18.4	3058	3	US-09-710-279-3585
6	229.5	18.4	3077	3	US-09-710-279-3594
7	225.5	18.1	729	3	US-09-134-001C-2317
8	214.5	17.2	768	3	US-09-107-532A-1266
9	210.5	16.9	1038	3	US-09-252-991A-11376

10	210.5	16.9	1086	3	US-09-252-991A-11418	Sequence 11418, A
c 11	204.5	16.4	11823	3	US-08-956-171B-136	Sequence 136, App
c 12	204.5	16.4	11823	3	US-08-781-986A-136	Sequence 136, App
13	199.5	16.0	792	3	US-09-328-352-360	Sequence 360, App
14	197	15.8	753	3	US-09-471-803A-2	Sequence 2, Appli
15	197	15.8	753	3	US-10-608-504-2	Sequence 2, Appli
16	197	15.8	876	4	US-09-605-703B-1839	Sequence 1839, Ap
17	197	15.8	1271	3	US-09-471-803A-1	Sequence 1, Appli
18	197	15.8	1271	3	US-09-471-803A-6	Sequence 6, Appli
19	197	15.8	1271	3	US-10-608-504-1	Sequence 1, Appli
20	197	15.8	1271	3	US-10-608-504-6	Sequence 6, Appli
21	194.5	15.6	621	3	US-09-134-000C-3384	Sequence 3384, Ap
22	187	15.0	675	3	US-09-583-110-2173	Sequence 2173, Ap
c 23	184	14.7	4792	3	US-08-961-527-156	Sequence 156, App
c 24	183	14.7	834	4	US-09-605-703B-2341	Sequence 2341, Ap
25	171	13.7	555	3	US-09-107-433-738	Sequence 738, App
26	169	13.5	588	3	US-09-902-540-2152	Sequence 2152, Ap
27	169	13.5	758	3	US-09-902-540-1415	Sequence 1415, Ap
28	146.5	11.7	693	3	US-09-543-681A-1	Sequence 1, Appli
29	143	11.5	867	3	US-09-489-039A-5047	Sequence 5047, Ap
c 30	136	10.9	1830121	3	US-09-557-884-1	Sequence 1, Appli
c 31	136	10.9	1830121	3	US-09-643-990A-1	Sequence 1, Appli
c 32	136	10.9	1830121	3	US-10-158-865-1	Sequence 1, Appli
c 33	124.5	10.0	3793	3	US-09-902-540-587	Sequence 587, App
34	122.5	9.8	744	3	US-09-328-352-1898	Sequence 1898, Ap
35	113.5	9.1	609	3	US-09-902-540-6892	Sequence 6892, Ap
36	103.5	8.3	16075	3	US-09-096-942-1	Sequence 1, Appli
37	103.5	8.3	16075	3	US-09-096-867-1	Sequence 1, Appli
38	103.5	8.3	1664976	3	US-08-916-421B-1	Sequence 1, Appli
39	103.5	8.3	1664976	3	US-09-692-570-1	Sequence 1, Appli
40	103	8.3	1647	3	US-09-328-352-2479	Sequence 2479, Ap
41	102.5	8.2	18355	3	US-08-956-171B-67	Sequence 67, Appl
42	102.5	8.2	18355	3	US-08-781-986A-67	Sequence 67, Appl
43	102	8.2	921	3	US-09-328-352-1272	Sequence 1272, Ap
44	101.5	8.1	1545	4	US-09-605-703B-2801	Sequence 2801, Ap
45	101.5	8.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-489-039A-3016
; Sequence 3016, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ IDS NOS: 14342
; SEQ ID NO 3016
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3016

Alignment Scores:
Pred. No.: 5.99e-116 Length: 756
Score: 1013.50 Matches: 200
Percent Similarity: 88.3% Conservative: 18
Best Local Similarity: 81.0% Mismatches: 26
Query Match: 81.2% Indels: 3
DB: 3 Gaps: 1

US-10-073-293A-4 (1-245) x US-09-489-039A-3016 (1-756)

Qy 1 MetGluSerProThrProGlnProAla-----ProGlySerAlaThrPheMetGlu 17
|||||:|||||
Db 10 ATGGAAGACCCCGCGCCCTTAACTGTGCTTACCTGACGCGCTGCCACCGTGGCGAA 69
|||||

Qy 18 GlyCysLysAspSerLeuProIleValIleSerThrValProValAlaPheAlaPheGly 37
 Db 70 GCGCTCAAGACAGCTCTCCCATGTCATCAGCTACCTTCGGGTGGCATTCGCTTCGGA 129
 Qy 38 LeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIle 57
 Db 130 CTCAATGCTACCCGCTGGGCTTACCCCGCTGGAAGCCCTGTTCTTCTCTGCTATATT 189
 Qy 58 TyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrp 77
 Db 190 TAGCCGCGCGCAGCAGTTGTGATCACCGCCATGTTGGCCGCGCAGTTGCTGTGG 249
 Qy 78 IleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuThrGlyProSerLeu 97
 Db 250 GTCGCGCGCTTACCGGTGATGCGATGTCGCGCACGCTGCTGATGGCCCTTCACTG 309
 Qy 98 ArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeu 117
 Db 310 CGCAGCCGCATTCGCTCAGCGCTGATGATGATGATGATGATGATGATGATGATGATG 369
 Qy 118 ThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgTrpSer 137
 Db 370 ACCGATGAGGTTTTCGCGCGCTTACCGCGCAGGCTGTTGCGGATTAACCGCGCTGGAGC 429
 Qy 138 GluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrVal 157
 Db 430 GAAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 489
 Qy 158 IleGlyAlaPheSerGlySerGlyLeuGlnGlyTyrProAlaValGluAlaAlaLeu 177
 Db 490 GCCGGCGCTGGTCCGGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
 Qy 178 GlyPheMetLeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgGlyGln 197
 Db 550 GGCTTTATGCTGCGCGCGCTGTTATGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
 Qy 198 SerLeuCysValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIle 217
 Db 610 TCACTCTGCTGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCTGCTGCT 669
 Qy 218 ProValAlaIleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPhe 237
 Db 670 CCGCGCGCGATCTCCCGCGAATGCTGCGCGGCTGCTGCGCGGCTGCTGCGCGGCTGCT 729
 Qy 238 TrpGlnGlyAlaProAspGlu 244
 Db 730 CTCGAGGGGATGCGCGATGAA 750

RESULT 2

US-09-328-352-2399
 ; Sequence 2399, Application US/09328352
 ; Patent No. 6562958

GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 2399

LENGTH: 801
 TYPE: DNA

ORGANISM: Acinetobacter baumannii
 US-09-328-352-2399

Alignment Scores:
 Pred. No.: 1.29e-63 Length: 801

Score: 592.00 Matches: 111
 Percent Similarity: 71.0% Conservative: 53

Best Local Similarity: 48.1% Mismatches: 67
 Query Match: 47.4% Indels: 0

DB: 3 Gaps: 0

US-10-073-293A-4 (1-245) x US-09-328-352-2399 (1-801)

Qy 7 GlnProAlaProGlySerAlaThrPheMetGluGlyCysLysAspSerLeuProIleVal 26
 Db 76 CAGATGGCTGCTCAACACAGCACTTTTGGCAAGTGCCAAAGATAGTCAAGCCATCGTT 135
 Qy 27 IleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSer 46
 Db 136 CTCACCTACTTACCACTTCTTTTGGCTTTGCGCTTCTGCACTCATCGTTTGGTTTACA 195
 Qy 47 ProLeuGluSerValPhePheSerCysIleThrAlaGlyAlaSerGlnPheValIle 66
 Db 196 CCATGGGAGCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 255
 Qy 67 ThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaLeuThrValMetAlaMet 86
 Db 256 GTGCGCACTTTTAGCAAGTGGTCTTCTATATATGTTAAACCGCTTTAACAGTCAATGCACTT 315
 Qy 87 AspValArgHisValLeuThrGlyProSerLeuArgSerArgIleIleGlnArgLeuGln 106
 Db 316 GATATACGACATGTAATTTATGGCCCTGCTCTTATATATTTAAATTTCCCACTTAATTAAT 375
 Qy 107 LysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGluValPheAlaAlaAlaThr 126
 Db 376 TTAAGAAGACAGCCGTTTGGGCTTGGGTTTAAAGACGAAAGTTTTCAGTGGTATG 435
 Qy 127 AlaLysLeuValArgAsnAsnArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPhe 146
 Db 436 ATCCAACCTTTCACAGCGCAGACAACTGGTCCGAGTCTTGATGTTGGGCTTAAGTCTA 495
 Qy 147 SerSerTrpSerSerTrpValPheGlyThrValIleGlyAlaPheSerGlySerGlyLeu 166
 Db 496 TTTAGCTGGATGCTTGGGCAAGTGGTTCATATTATAGCGGGTATTATTCAGATCAAGTC 555
 Qy 167 LeuGlnGlyTyrProAlaValGluAlaLeuGlyPheMetLeuProAlaLeuPheMet 186
 Db 556 GCACACCTTCCAAAGTTTATACAGCAGCTTTAGATTCTTATTACCTGCTTTATTTTA 615
 Qy 187 SerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCysValThrAlaAlaLeuVal 206
 Db 616 AGTTTTTATTAGCGCATTTGAACGAAAGCATAGCTTGTAGTTGCTGTATCGTTAGTG 675
 Qy 207 GlyAlaLeuAlaGlyValThrLeuPheSerIleProValAlaIleLeuAlaGlyIleVal 226
 Db 676 GTGTCAGCTTTGGCATGTTATTGGATTAACTATCAGCGGCTATCTTTATTGGCATCTCT 735
 Qy 227 CysGlyCysLeuThrAlaLeuIleGlnAlaPhe 237
 Db 736 TCAGGCACTTTAGCAGGCTCTATTAAATATTAT 768

RESULT 3

US-09-489-039A-1307

; Sequence 1307, Application US/09489039A

; Patent No. 6610836

GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-27

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 1307

; LENGTH: 744

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-1307

Alignment Scores:

Pred. No.: 8.4e-21 Length: 744
Score: 246.50 Matches: 67
Percent Similarity: 48.3% Conservative: 46
Best Local Similarity: 28.6% Mismatches: 106
Query Match: 19.8% Indels: 15
DB: 3 Gaps: 6

US-10-073-293A-4 (1-245) x US-09-489-039A-1307 (1-744)

Qy 18 GlyCysLeuSerLeuProLeuValleValleSerTyrIleProValAlaPheAlaPheGly 37
Db 28 GCGCCAGCGTATATGCCCTATGCTATGCGGACCTTCCATTTTCTATTTATGTCGCG 87
Qy 38 LeuAlaAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIle 57
Db 88 GCATTAAGCGTAGCGCGGATGAGCGTCTGGCAGAGTACGGCGTGGCGGATCGTC 147
Qy 58 TyrAlaGlyAlaSerGlnPheValleThrAlaMetLeuAlaAlaGlySerSerLeuTrp 77
Db 148 ATTGCGGCTCGCGCAATGCTGGCAGTGAATATGTTTAAACCGCGGCGCCAGCTCGC 207
Qy 78 IleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeu 97
Db 208 GTGATTATTTACCAACATTAATATCAATCTTCCACGCTGCTATAGCGGCTCTATT 267
Qy 98 ArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeu 117
Db 268 ---TCCGCGACCGTGGCGAAGCTTCATCTTTTAAATAATGCTTTATGCTCTATGCGCTC 324
Qy 118 ThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTrpSer 137
Db 325 ACCGACGAGGTGTAT---GCCACACCGCTCAAGAGATGGGAAGAAATATAAAGGAAAA 381
Qy 138 GluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerTrpValPheGlyThrVal 157
Db 382 TATCTGTTTATGTTGGCGATGATACCTTCTGGCGGATTTGGGCTGGCTGATTTT 441
Qy 158 IleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAla---Ala 176
Db 442 CTTCGCGCGCTGGTGGCGCC-----TCGTTTCCCACTGAGAAATACGGT 489
Qy 177 LeuGlyPheMetLeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLys 196
Db 490 CTCGATTTGCGCATGGTGGCGGCTTTATCGCCATCGTGGTGGCG-----CAAATATA 543
Qy 197 GlnSerLeuCysValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSer 216
Db 544 AGCAGGCTGACCGTGGCGGCTGTATGACCCCGGTTTCCGCGGCTGCTGTAGTGGTC 603
Qy 217 IlePro-----ValAlaIleLeuAlaGlyIleValCysGlyCysLeu 230
Db 604 CTACCTATTGCTGGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
Qy 231 ThrAlaLeuIleGlnAlaPheTrpGlnGlyAlaProAspGlu 244
Db 664 GTCGACCTGGCGGAAGAGCGAAGCAGATGTCGCAAAACCGAA 705

RESULT 4

US-09-710-279-267
; Sequence 267, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 267
; LENGTH: 693

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-267

Alignment Scores:

Pred. No.: 9.66e-19 Length: 693
Score: 229.50 Matches: 65
Percent Similarity: 49.3% Conservative: 47
Best Local Similarity: 28.6% Mismatches: 102
Query Match: 18.4% Indels: 13
DB: 3 Gaps: 6

US-10-073-293A-4 (1-245) x US-09-710-279-267 (1-693)

Qy 14 ThrPheMetGluGlyCysLysAspSerLeuProIleValleSerTyrIleProValAla 33
Db 13 ACGTTTAAACAGAGGTGTAAAGAGAGTATTCCACAGTTACTTGGTTATGCGAGGTGAGGA 72
Qy 34 PheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePhe 53
Db 73 CTATCGTGTGGAATTTGGCAGCTCTCCAAAATTTTCAAGTTTGTAGAAATATTATTATTG 132
Qy 54 SerCysIleIleTyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGly 73
Db 133 TGCTGATTTATGCTGCTGCGAGCTCAATTTATTATTGTCATTTAGTATGATGTCAGGC 192
Qy 74 SerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyr 93
Db 193 ACCCTATTTCTGCAATTTGCTTACAAATCTTATCGTAACTCTCGAATGTTCTTATTA 252
Qy 94 GlyProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrp 113
Db 253 AGTATGACTTTAGCACCTAATTAAGCAATATGGAATTTGGATAGGCTGGCTTGGGA 312
Qy 114 AlaPheGlyLeuThrAspGluValPheAlaAlaThrAlaLysLeuValArgAsnAsn 133
Db 313 ACGTTA---TTACAGATGAACTTTTGGCGTTGCTATAACACCATATGTTAAA---GGT 366
Qy 134 ArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerTrpVal 153
Db 367 GAAATAAATAACGATCGATCGCTACACGACTAAATATTAATGCTTACTTATTTTGGACT 426
Qy 154 PheGlyThrValIleGlyAlaPheSerGlySerGlyLeuGlnGlyTyrProAlaVal 173
Db 427 GTTTCCTGTGTAATCGGTGCCATTTTTCGGAGAG-----TATATTTCAAT 471
Qy 174 GluAlaAlaLeuGly-----PheMetLeuProAlaLeuPheMetSerPheLeuLeuAla 191
Db 472 CCTGATGCGCTCGGCTAGACTTTTGCATTTACCGCAATGTTATTTTATGTTATATCT 531
Qy 192 SerPheGln-----ArgLysGlnSerLeuCysValThrAlaAlaLeuValGlyAlaLeu 209
Db 532 CAATTTGAAGGATTAAGAAATCAGATTCAGAAATATATATGTTACTTATGTTATGTTGTG 591
Qy 210 AlaGlyValThrLeuPhe-----SerIleProValAlaIleLeuAlaGlyIleValCys 227
Db 592 ATTTGATGATGCTTCTTCAAGTTCATTTCACTTACCTTACCTACCTAGCAATTTTAATAGCC 651
Qy 228 GlyCysLeuThrAlaLeuLeu 234
Db 652 GCAATGTTGCTGCATTTGTTA 672

RESULT 5

US-09-710-279-3585/c
; Sequence 3585, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US

Qy 210 AlaGlyValThrLeuPhe-----SerIleProValAlaIleLeuAlaGlyIleValCys 227
Db 2975 ATTGGATGATGCTCTCTTAAGTTCAATTCATCTTACCTTCACCTAGCAATTTTAATAGCC 3034
Qy 228 GlyCysLeuThrAlaLeuIle 234
Db 3035 GCAATTGTTGTCGATTTGTTA 3055

RESULT 7

US-09-134-001C-2317
; Sequence 2317, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2317
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2317

Alignment Scores:

Pred. No.: 3,28e-18 Length: 729
Score: 225.50 Matches: 66
Percent Similarity: 45.5% Conservative: 39
Best Local Similarity: 28.6% Mismatches: 89
Query Match: 18.1% Indels: 37
DB: 3 Gaps: 7

US-10-073-293a-4 (1-245) x US-09-134-001C-2317 (1-729)

Qy 14 ThrPheMetGluGlyCysLysAspSerLeuProIleValIleSerTyIleProValAla 33
Db 49 ACGTTTAAACAAAGTGTGAAGAGTGTATTCACAGTACTTGTGTTATGCTGGGTAGGA 108
Qy 34 PheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGlySerValPhePhe 53
Db 109 CTATCGTTGGAAATGTGGAGCTCCCAAAATTCAGTGTTTAGAAATATTTTATTG 168
Qy 54 SerCysIleIleTyAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGly 73
Db 169 TGTCTGATTTATTTATGCTGGTGCAGCCCAATTTATTTTGTACATTAGTAGTGCAGGC 228
Qy 74 SerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTy 93
Db 229 ACCCTATTTCGAATTTGTCTTACACACTTATGTTACTCTCGAATGTTCTTATTA 288
Qy 94 GlyProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrp 113
Db 289 AGTATGACTTTAGCACCAATTAAGCAATATGATTTTGGAAATGGGTAGGCGCTTGA 348
Qy 114 AlaPheGlyLeuThrAspGluValPheAlaAlaThrAlaLysLeuValArgAsnAsn 133
Db 349 ACGTTA---TTAACAGATGAACCTTTTGGAGTGTCTATAACGCCATATGTCAAA--GGT 402
Qy 134 ArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerSerTrpVal 153
Db 403 GAAAAAATTAACGATCGATCGTCTACACGCACTAAATATTAATCTGCTTACTTTTGGACT 462
Qy 154 PheGlyThrValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyProAlaVal 173
Db 463 GTTTCATGTGTAATTCGGTGCATTTTCGGAGAG-----TATATTTCAAA 507
Qy 174 GluAlaAlaLeuGly-----PheMetLeuProAlaLeuPheMetSerPheLeuLeuAla 191

Db 508 CCTGATCGCTGGCCTAGACTTTTGCCATTACCGCAATGTTTATTTTATGATATATCT 567
Qy 192 SerPheGln-----ArgLysGlnSerLeu-----CysVal 201
Db 568 CAATTTCAGGGATTAGAAATACAGATTCAGATATATATTTGTAATCTGATGTTG 627
Qy 202 -----ThrAla 203
Db 628 ATTGTGATGATGCTTCTTCTAAGTTCAATTCATCTTACATGACGATGCAATTTTAATAGCC 687
Qy 204 AlaLeuValGlyAlaLeuAlaGlyValThrLeu 214
Db 688 GCAATTGTTGCTGCAATTTAGGGTGTGTGATG 720

RESULT 8

US-09-107-532A-1266
; Sequence 1266, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; TELEFAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 1266:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...768
; SEQUENCE DESCRIPTION: SEQ ID NO: 1266:

US-09-107-532A-1266

Alignment Scores:
Pred. No.: 8.24e-17 Length: 768
Score: 214.50 Matches: 66
Percent Similarity: 46.8% Conservative: 43
Best Local Similarity: 28.3% Mismatches: 101

Query Match: 17.2% Indels: 23
DB: 3 Gaps: 6

US-10-073-293A-4 (1-245) x US-09-107-532A-1266 (1-768)

QY 20 LysAspSerLeuProIleValIleSerTyrlleProValAlaPheAlaPheGlyLeuAsn 39
DB 94 AAAGACAGTTACCTACCGTTTCGGTTATATCGGTATGGACTTGCAATTGGTATCGTT 153

QY 40 AlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrla 59
DB 154 GGGAAAGCTGCGGATTTTCATCCACTAGTGTGTCACGTTGATGTCCTGCTGCTATGCT 213

QY 60 GlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAla 79
DB 214 GGTTCGCCCAATTTATCAGATGCAATGCTGTCTAGTCACAGCCCAATTTCCATC 273

QY 80 AlaLeuThrValMetAlaMetAspValArgHisValLeuTyrlGlyProSerLeu----- 97
DB 274 GTTTTCGACCTTCTAGTCAATTCGCCGAATGATCTGATGAGTATGACGATGCTCCT 333

QY 98 -----ArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAla 114
DB 334 TATTTCAAGAAATCGGCTGCTCAAAATCTG-----CTGATTGGC 375

QY 115 PheGlyLeuThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArg 134
DB 376 ACATTGCTGACAGATGAAAGTTTCGACCTGCGAATGAATGAATGAATGAACTACTGGA 435

QY 135 ArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerTrpValPhe 154
DB 436 AAGTTAACTTCGTGGATGAATACACCTAATGGATTTCTATTGACGTGGTGGGT 495

QY 155 GlyThrValIleGlyAlaPheSerGlySerGlyLeuGlnGlyTyrlProAlaValGlu 174
DB 496 TCTTCTCTGTGGGCACTTTAGGAAATTTTATTACTGAT-----CCGAAAAAG 546

QY 175 AlaAlaLeuGlyPheMetLeuProAlaLeuPheMetSer-----PheLeuLeuAlaSer 192
DB 547 TTTGGTTGGATTTTCGCCATGCTGCTATGTTTATCGGGTTACTTTATCTTCAAGTCA 606

QY 193 PheGlnArgLysGlnSerLeuCysValThrAlaAlaLeuValGly-----AlaLeu 209
DB 607 TCTGATGAACACCTCTAAAGCATACAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTG 666

QY 210 AlaGlyValThrLeuPheSerIlePro-----ValAlaIleLeuAlaGlyIle 225
DB 667 GTATATGTCGGACTGATTTTATCCCAAGCAACTTAGTGATTTAGTTGTTAACAACATG 726

QY 226 ValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrp 238
DB 727 GGGTGGGATTTGGGGGTGGGATCAACATGCTTCTTT 765

RESULT 9

US-09-252-991A-11376
; Sequence 11376, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11376
; LENGTH: 1038
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11376

Alignment Scores:
Pred. No.: 4.15e-16 Length: 1038
Score: 210.50 Matches: 73
Percent Similarity: 44.3% Conservative: 43
Best Local Similarity: 27.9% Mismatches: 108
Query Match: 16.9% Indels: 38
DB: 3 Gaps: 11

US-10-073-293A-4 (1-245) x US-09-252-991A-11376 (1-1038)

QY 1 MetGluSerProThrProGlnPro-AlaProGlySerAla-----Th 14
DB 297 CTGACGCGAGCTTCACCAACCGCCATGCCCATGACATGCCACGACCTGCCCCCAAGC 356

QY 14 rPheMetGluGlyCysLysAspSerLeuProIleValIleSerTyrlleProValAlaPh 34
DB 357 CTTTCTCGGGCGCGCTGCGCATCTCCCATTTGCTGCGCGCTGCGCGCTGCGGGCT 416

QY 34 eAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSe 54
DB 417 GCTGCGGCGATCGATGCGCCATCGAAGCCAACTCACCCTCGGCGCCGAGGCGCCAGGCGCTGTC 476

QY 54 rCysIleIleTyrlAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySe 74
DB 477 CGCATGCTCTTCGCGCGCGCGCCAGTTGGTGGCATGCGCATGCTCAAGGCGGGTGC 536

QY 74 rSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyrlG 94
DB 537 CGGCTTCTTTTCGATCATCTTACACCTCTCTGACCTCTGACCTCGCAGCACCTGTCTACGG 596

QY 94 yProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAl 114
DB 597 CATGAGCTTGGGCGCG---GTCTGTCACTCTG-----CCGGAGACGCTGCGC 641

QY 114 aPheGly-----LeuThrAspGluValPheAlaAlaAlaThrAlaLysLeuVa 130
DB 642 GATCGCGCTGGGCTTCTGCTCACCGACGAGTTCTTTCGCCCTCGCCAGCAGCAGCCG 701

QY 130 lArgAsnAsnArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSe 150
DB 702 GCGCAACTTCAACCGCTG-----TAGCCCTCGCGCTGCTGCTGACCTTCTATAT 752

QY 150 rSerTrpValPheGlyThrValIleGlyAlaPheSerGlySerGlyLeuGlnGlyTy 170
DB 753 GGCCTGGAACCTGTTACCTTCGCGCGCATCTCTGCGCGCAGCATT-----801

QY 170 rProAlaValGlu---AlaAlaLeuGlyPheMetLeuProAlaLeuPheMetSerPheLe 189
DB 802 -CCGGTCTCGAACACCTGGGCGCTGGATTCTCCATCGCCGCCACCTTCATCGCGCTGT 860

QY 189 uLeuAlaSerPheGlnArgLysGlnSerLeu---CysValThrAlaAlaLeuValGlyAl 208
DB 861 CGCGCGCTGGTGGCAACAGTACACACTGCTGCTGCTGCGCGCACCTGCTGCTGCTGCTC 920

QY 208 aLeuAlaGlyValThrLeuPheSer-----IleProValAlaAl 221
DB 921 GGTG-----CTGTTTCAGCCACTGGCAATGGAGTTTCGGCGCTGCTGCTGCGCGG 968

QY 221 eLeuAlaGlyIleValCysGly-----CysLeuThrAlaLeuIleGlnAlaPheTrpGl 239
DB 969 TCTCGCGGATGCGCGCGGTTTCACTGCAACAAGCTGTATCGGAGGCGCCCATGCTC 1028

QY 239 nGly 240
DB 1029 TGGG 1032

RESULT 10

US-09-252-991A-11418
; Sequence 11418, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

[illegible]

```

Qy 34 PheAlaPheGlyLeuAenAlaThrArgLeuGlyPheSerProLeuGluSerValPhePhe 53
Db 1166 ATTCAITGGTATGGCTTCGCTCAAACTTTAGTATTTAGAAATTTGCTTGTGA 1107
Qy 54 SerCysIleIleTyAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGly 73
Db 1106 TGTCTTGTTATATGCGGCTGCTGCGCAATTTATATGTCGCGTTGTTTATAGCAGT 1047
Qy 74 SerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTy 93
Db 1046 ACACCTATATCAGCGATTTGTAATCTATCTGTAATTTATCTGTAATTTATCTCTTTA 987
Qy 94 GlyProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThr--AlaLeu 112
Db 986 AGTAGTGCCTTCGCCAAACTTC-----AAGACATATGGGGTTT 948
Qy 113 Trp---AlaPheGlyLeu-----ThrAspGluValPheAlaAlaAlaThrAla 127
Db 947 TGGAAACCGTGTGGATTAGGTTTCATTAGTAATGACGAAACGTTTGGCGTCGCCATTACA 888
Qy 128 LysLeuValArgAsnAsnArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSer 147
Db 887 CCTTATTTAAAA---GGAGAAAGCTATCAATGATCGTTGCATGCTGTTTAAACATCACA 831
Qy 148 SerTrpSerSerTrpValPheGlyThrValIleGlyAlaPheSerGlySerGlyLeuLeu 167
Db 830 GCATATTTATTTTGGGCAATTTTCATGTGTAGCTGGCGCTTTTATTGGCGAA----- 780
Qy 168 GlnGlyTyProAlaValGluAlaLeuGly-----PheMetLeuProAlaLeuPhe 185
Db 779 -----TATATCTCAATTCGCAACGCTAGGTTAGATTTTGTATCAGCGCTATGTTT 726
Qy 186 MetSerPheLeuLeuAlaSerPheGln-----ArgLysGlnSerLeuCysValThrAla 203
Db 725 ATCTTTTGGCCATGCGCAATTTGAATCAATTAATCGGATTAATCGGATTAAGAAATTACATA 666
Qy 204 AlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProVal-----Ala 220
Db 665 GTACTCATTTATGCGTCATATA---GTAATGATTTATCGCTAAGTATGTTTATGCTCTTCA 609
Qy 221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIle 234
Db 608 TATCTAGCAATATTAATTTGCAGCCACAAATTTTCAGCAGCGTTA 567

```

RESULT 12

```

US-08-781-986A-136/c
; Sequence 136, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 136:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11823 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-136

```

```

Alignment Scores:
Pred. No.: 1.06e-13 Length: 11823
Score: 204.50 Matches: 65
Percent Similarity: 47.0% Conservative: 45
Best Local Similarity: 27.8% Mismatches: 97
Query Match: 16.4% Indels: 27
DB: 3 Gaps: 10

```

US-10-073-293A-4 (1-245) x US-08-781-986A-136 (1-11823)

```

Qy 14 ThrPheMetGluGlyCysLysAspSerLeuProIleValIleSerTyIleProValAla 33
Db 1226 AGTTTATAGCAAGCGTCGACAGATGTATCCACATATTATGGGTATGCCGTTGGT 1167
Qy 34 PheAlaPheGlyLeuAenAlaThrArgLeuGlyPheSerProLeuGluSerValPhePhe 53
Db 1166 ATTTCAITGGTATTTGGCTTCGCTCAAACTTTAGTATTTAGAAATTTGCTTGTGA 1107
Qy 54 SerCysIleIleTyAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGly 73
Db 1106 TGTCTTGTTATATGCGGCTGCGCAATTTATATGTCGCGTTTATTATGACAGT 1047
Qy 74 SerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTy 93
Db 1046 ACACCTATATCAGCGATTTGTAATCTATCTGTAATTTATCTGTAATTTATCTCTTTA 987
Qy 94 GlyProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThr--AlaLeu 112
Db 986 AGTAGTGCCTTCGCCAAACTTC-----AAGACATATGGGGTTT 948
Qy 113 Trp---AlaPheGlyLeu-----ThrAspGluValPheAlaAlaAlaThrAla 127
Db 947 TGGAAACCGTGTGGATTAGGTTTCATTAGTAATGACGAAACGTTTGGCGTCGCCATTACA 888
Qy 128 LysLeuValArgAsnAsnArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSer 147
Db 887 CCTTATTTAAAA---GGAGAAAGCTATCAATGATCGTTGCATGCTGTTTAAACATCACA 831
Qy 148 SerTrpSerSerTrpValPheGlyThrValIleGlyAlaPheSerGlySerGlyLeuLeu 167
Db 830 GCATATTTATTTTGGGCAATTTTCATGTGTAGCTGGCGCTTTTATTGGCGAA----- 780
Qy 168 GlnGlyTyProAlaValGluAlaLeuGly-----PheMetLeuProAlaLeuPhe 185
Db 779 -----TATATCTCAATTCGCAACGCTAGGTTAGATTTTGTCTATCAGCGCTATGTTT 726
Qy 186 MetSerPheLeuLeuAlaSerPheGln-----ArgLysGlnSerLeuCysValThrAla 203
Db 725 ATCTTTTGGCCATGCGCAATTTGAATCAATTAATCGGATTAATCGGATTAAGAAATTACATA 666
Qy 204 AlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerIleProVal-----Ala 220
Db 665 GTACTCATTTATGCGTCATATA---GTAATGATTTATCGCTAAGTATGTTTATGCTCTTCA 609
Qy 221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIle 234
Db 608 TATCTAGCAATATTAATTTGCAGCCACAAATTTTCAGCAGCGTTA 567

```

RESULT 13

US-09-328-352-360
; Sequence 360, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 360
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-360
Alignment Scores:
Pred. No.: 6.28e-15 Length: 792
Score: 199.50 Matches: 65
Percent Similarity: 47.3% Conservative: 41
Best Local Similarity: 29.0% Mismatches: 101
Query Match: 16.0% Indels: 17
DB: 3 Gaps: 6
US-10-073-293A-4 (1-245) x US-09-328-352-360 (1-792)
Qy 15 PheMetGluGlyCysLysAspSerLeuProIleValIleSerTyrIleProValAlaPhe 34
Db 133 TTTTACGAGGGCGGATAGATATATGCGCATCTATCTATTTCCGTTATACCATGGCGTATT 192
Qy 35 AlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSer 54
Db 193 TTAGCTGGATCAATGAGTTCATGCGGTTTATCATCTTTTATAAGCACTAGCGCATGCA 252
Qy 55 CysIleIleTyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAAGlySer 74
Db 253 GGAATTGTTGTTGCGCGCTGCGCAATTAGTAAGTTTAAAGCATGGTGATGGAAGGGGCA 312
Qy 75 SerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHISValLeuTyrGly 94
Db 313 TCGTTATTAACCATCTATGTCACCTATTTCTTCTTCTGCTCAACATTTTATTATGCT 372
Qy 95 ProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAla 114
Db 373 TTAACCTTAAGAAATGAC---ATTCTATTGCTCTCTCATAAAGATTAACCTTAGGT 429
Qy 115 PheGlyLeuThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnArg 134
Db 430 TTTCTGCTTACCGAGAAATTATTTGCGAGTAAGTGTCTCT-----AATGAG 474
Qy 135 ArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerSerTrpValPhe 154
Db 475 AAAAACAATCTCAATATCTATTCGAGCAGGTCTTGTCTTTTATTATTATTTGGGTGTT 534
Qy 155 GlyThrValIleGlyAlaPheSerGlySerGlyLeu-----LeuGlnGlyTyrProAla 172
Db 535 TTTAGTTTAGTGTGTTATTTTATTAGCAACGCGCTACCTAACTATTATAATTAT- 588
Qy 173 ValGluAlaAlaLeuGlyPheMetLeuProAlaLeuPheMetSerPheLeuLeuAlaSer 192
Db 589 -----CATTTGGATTTTTCATTTATGCTATTTTGTGTCATGATTTGTTCCVATG 639
Qy 193 PheGlnArgLysGlnSerLeuCysValThrAlaAlaLeuValGlyAlaLeuAlaGlyVal 212
Db 640 TGTAAAGGAGCGCT-----GTAAATGGCAGGCAATTTTAATGCTGTGTCAGTGGTTT 693
Qy 213 ThrLeu-----PheSerIleProValAlaIleLeuAlaGlyIleValCysGlyCysLeu 230
Db 694 GTTCTTAAATTTTTCATATAGAGGTGCCATTCTAATTTCTGTTTATTGGGGATGTTT 753
Qy 231 ThrAlaLeuIle 234
||||:||||

Db 754 ATTCGGGTAATA 765
RESULT 14
US-09-471-803A-2
; Sequence 2, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELE, LOTHAR
; APPLICANT: PEPPERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471.803A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(753)
; OTHER INFORMATION: brnF
; OTHER INFORMATION: ATCC14752
US-09-471-803A-2
Alignment Scores:
Pred. No.: 1.18e-14 Length: 753
Score: 197.00 Matches: 68
Percent Similarity: 46.6% Conservative: 41
Best Local Similarity: 29.1% Mismatches: 95
Query Match: 15.8% Indels: 30
DB: 3 Gaps: 9
US-10-073-293A-4 (1-245) x US-09-471-803A-2 (1-753)
Qy 17 GluGlyCysLysAspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPhe 36
Db 100 CAAGGTCTAAACCTCCCTGCTGCGAGTTTGGCATGTACCGATTGGTATTGCGTTT 159
Qy 37 GlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIle 56
Db 160 GGTCTCTTGTGTTATTCAATACGGCTACGAATGGTGGCGACCCCACTGTTTTCGGCGCTG 219
Qy 57 IleTyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeu 76
Db 220 ATTTTCGGGGCTCCACCGAAATGCTGTCATCGCTCTGTTGGGGCGACGCCCTG 279
Qy 77 TrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSer 96
Db 280 GCGCGCATCGCTCACCACATTGCTGGTGAACCTTCGCCACGATTCTATCGCTTTCA 339
Qy 97 LeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrp-----Ala 114
Db 340 TTCCCGCTGTCATGTGTC-----AAAAACCCCATTCGCCCTTCTATTTCGTT 387
Qy 115 PheGlyLeuThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnArg 134
Db 388 TTCCGCTTATCCAGCAAGCTACGCACTACGCGGCTGCGCGCAGG-----CCGCA 435
Qy 135 ArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerSerTrpValPhe 154
Db 436 GGCTGGTGGCGGGCGGACTTATCTCAATGCAAAAGCGTTTCACTCTACTGGGTATTC 495
Qy 155 GlyThrValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGlu 174
Db 496 -----GGCGGTCTTCACCGAGTGGCGATGCGAGTTGATTCTCTTTTGAATT 543


```

Qy 175 AlaAlaLeuGlyPheMetLeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGln 194
Db 544 AAGGGCCCTCGAGTTGGCCCTTGTCTCTTGTGCAGCTGAGCTTTGGATTCCTGCCGA 603
Qy 195 ArgLysGln-----SerLeuCysValThrAlaAlaLeuVal 206
Db 604 ACGAAAAGCAGATCCCTTCTGCTGCTCGCAGGTTTGAGCTTCACCATTCCTTGTG 663
Qy 207 GlyAlaLeuAlaGlyValThrLeuPheSerIleProValAlaIleLeuAlaGlyIleVal 226
Db 664 ---GTAATTCAGGTCAGGCCCTATTGCG-----GGCTGCTG-----ATCTTC 705
Qy 227 CysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly 240
Db 706 TTGGGTCTGTTGACC-----ATCCGGTACTTCTTCTTGGGA 741

RESULT 15
US-10-608-504-2
; Sequence 2, Application US/10608504
; Patent No. 6841360
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELE, LOTHAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 7601/80525
; CURRENT APPLICATION NUMBER: US/10/608,504
; CURRENT FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: US 09/471,803
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(753)
; OTHER INFORMATION: brnF
; FEATURE:
; OTHER INFORMATION: ATCC14752
US-10-608-504-2

Alignment Scores:
Pred. No.: 1.18e-14 Length: 753
Score: 197.00 Matches: 68
Percent Similarity: 46.6% Conservative: 41
Best Local Similarity: 29.1% Mismatches: 95
Query Match: 15.6% Indels: 30
DB: 3 Gaps: 9

US-10-073-293A-4 (1-245) x US-10-608-504-2 (1-753)
Qy 17 GluGlyCysLysAspSerLeuProIleValIleSerTyIleProValAlaPheAlaPhe 36
Db 100 CAAGGCTCTAAACCTCCCTTGTCTGTCAGGTTTGGGCATGTACCCGATTTGGTATTCGCTTT 159
Qy 37 GlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIle 56
Db 160 GGTCTCTTGGTTATTCAATACGGCTACGAATGGTGGGCGAGCCCACTGTTTCCGGCCTG 219
Qy 57 IleTyAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaGlySerSerLeu 76
Db 220 ATTTTCGGGGCTCCACCGAAATGTGTCTCATCGCCCTCGTTTGGGGCGCAGCCCTG 279
Qy 77 TrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyTrpGlyProSer 96

```

```

Db 280 GGGCCCATCGCGCTCACCACATTCGTGTGAACCTTCGCCACGATTTCTATGCGTGTTC 339
Qy 97 LeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrp-----Ala 114
Db 340 TTCCCGCTGCATGTGTGTC-----AAAAACCCCATTTGCCGTTTCTATTTCGGTT 387
Qy 115 PheGlyLeuThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnArg 134
Db 388 TTCCCGCTTATCGACGAGCCTACGAGTCACTCGGCCAGG-----CCCGCA 435
Qy 135 ArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerSerTrpValPhe 154
Db 436 GGCTGTGCGCGTGGCGACTTATCTCAATGCAATAGGTTTCACTCCTACTGGGTATTC 495
Qy 155 GlyThrValIleGlyAlaPheSerGlySerGlyLeuGlnGlyTyProAlaValGlu 174
Db 496 -----GGCGGTCTCACCGAGTGGCGATCGCAGATTGATTCTCTTTGAAATT 543
Qy 175 AlaAlaLeuGlyPheMetLeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGln 194
Db 544 AAGGGCCTCGAGTTGGCCCTTGTCTCTTGTGCAGCTGAGCTTTGGATTCCTGCCGA 603
Qy 195 ArgLysGln-----SerLeuCysValThrAlaAlaLeuVal 206
Db 604 ACGAAAAGCAGATCCCTTCTGCTGCTCGCAGGTTTGAGCTTCACCATTCCTTGTG 663
Qy 207 GlyAlaLeuAlaGlyValThrLeuPheSerIleProValAlaIleLeuAlaGlyIleVal 226
Db 664 ---GTAATTCAGGTCAGGCCCTATTGCG-----GGCTGCTG-----ATCTTC 705
Qy 227 CysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly 240
Db 706 TTGGGTCTGTTGACC-----ATCCGGTACTTCTTCTTGGGA 741

```

Search completed: February 16, 2006, 19:06:12
Job time : 230.666 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:44:58 ; Search time 946.966 Seconds
(without alignments)
2139.461 Million cell updates/sec

Title: US-10-073-293A-4

Perfect score: 1248

Sequence: 1 MESPTQPAPGSATFMEGCK.....VCGCLTALIOAFMQGAPDEL 245

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10073293/runat_15022006_120740_1106/app_query.fasta_1
-DB=Published Applications NA Main -QMT=fastcap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p
-USER=US10073293 @CGN_1_1_2064 @runat_15022006_120740_1106 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA Main.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1248	100.0	738	9	US-10-073-293A-3
2	426	34.1	1589	3	US-09-764-877-739
3	426	34.1	1589	3	US-09-860-870-69
4	426	34.1	1589	3	US-09-764-875-55
5	426	34.1	1589	6	US-10-227-646-69
6	426	34.1	1589	6	US-10-242-515-739
7	426	34.1	1595	3	US-09-764-877-1016

Sequence 372, Appl
Sequence 1016, App
Sequence 242, App
Sequence 2949, Ap
Sequence 8, Appli
Sequence 2058, Ap
Sequence 3752, App
Sequence 136, App
Sequence 136, App
Sequence 289, App
Sequence 2, Appli
Sequence 1, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 1445, Ap
Sequence 4979, Ap
Sequence 2, Appli
Sequence 2597, Ap
Sequence 2599, Ap
Sequence 6651, Ap
Sequence 156, App
Sequence 156, App
Sequence 3408, Ap
Sequence 1, Appli
Sequence 3183, Ap
Sequence 1, Appli
Sequence 95, Appl
Sequence 1, Appli
Sequence 738, App
Sequence 4863, Ap
Sequence 4832, Ap
Sequence 2485, Ap
Sequence 572, App
Sequence 685, App
Sequence 1, Appli
Sequence 1, Appli
Sequence 54, Appl

ALIGNMENTS

RESULT 1

US-10-073-293A-3

Sequence 3, Application US/10073293A
Publication No. US20050239175A1
GENERAL INFORMATION:
APPLICANT: TABOLINA, EKATERINA
APPLICANT: RYBAK, KONSTANTIN
APPLICANT: KHOURGES, EVGENI
APPLICANT: VOROSHILOVA, ELVIRA
APPLICANT: GUSYATINER, MIKHAIL
TITLE OF INVENTION: METHOD FOR PRODUCING L-AMINO ACID USING BACTERIA BELONGING TO THE
FILE REFERENCE: 219594USO
CURRENT APPLICATION NUMBER: US/10/073,293A
CURRENT FILING DATE: 2002-02-13
PRIOR APPLICATION NUMBER: RU 2001103865
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: RU 2001104998
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: RU 2001104999
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: RU 2001117632
PRIOR FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: RU 2001117633
PRIOR FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 738
TYPE: DNA
ORGANISM: Escherichia coli

FEATURE:

; NAME/KEY: CDS
; LOCATION: (1) ..(735)
; OTHER INFORMATION:

US-10-073-293A-3

Alignment Scores:

Pred. No.: 1.18e-134 Length: 738
Score: 1248.00 Matches: 245
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-073-293A-4 (1-245) x US-10-073-293A-3 (1-738)

Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db 1 ATGGAAGCCCTACTCCACAGCCTGCTCTCGTTCGGCGACCTTCATGGAAGGATGCAAA 60
Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
Db 61 GACAGTTTACCGATTGTATTAGTTATATTCCGGTGGCCTTTGGCTTCGGTCTGAATGCG 120
Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
Db 121 ACCCGTCTGGGATTTCTCTCTCGAAAGCGTTTCTCTCGATCATTTATGCGAGGC 180
Qy 61 AlaSerGlnPheValIleThrAlaMetIleuAlaAlaGlySerSerLeuTrpIleAlaAla 80
Db 181 GCGAGCCAGTTCGTCAATTACCGCATGCTGGCAGCCGGAGTAGTTTGTGGATTGCTGCA 240
Qy 81 LeuThrValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArg 100
Db 241 CTGACCGTCATGCGAAATGAGTGTTCGCATGTTGTCATGCCCCGTCACTCGTAGCCGT 300
Qy 101 IleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGlu 120
Db 301 ATTATTACGCTCTGCAAAATCGAAACCGCCCTGTGGGGTTTGGCTGACGGATGAG 360
Qy 121 ValPheAlaAlaThrAlaLysLeuValArgAsnArgArgTTPSerGluAsnTrp 140
Db 361 GTTTTTCGCGCGCAACCGCAAACTGGTAGCGCAATAATCGCCGCTGGAGCGAGAACTGG 420
Qy 141 MetIleGlyIleAlaPheSerSerTrpSerSerTrpValPheGlyThrValIleGlyAla 160
Db 421 ATGATCGCATTCCTTCAGTTTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 480
Qy 161 PheSerGlySerGlyLeuLeuGlnGlyTyrProAlaValGluAlaAlaLeuGlyPheMet 180
Db 481 TTCTCCGGCAGCGCTTGTCTCAAGTTATCCCGCGTTGAAGCTGCATTAAGTTTATG 540
Qy 181 LeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLeuCys 200
Db 541 CTTCGGGCACCTTTATGAGTTCTCTGTCGCCCTTTCCAGCGCAACAATCTCTTTGC 600
Qy 201 ValThrAlaAlaLeuValGlyAlaIleuAlaGlyValThrLeuPheSerIleProValAla 220
Db 601 GTTACCGCAGCGTTAGTTGGTGGCCCTTCGAGCGGTAAACGCTATTTCCTATTCGCGCC 660
Qy 221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly 240
Db 661 ATTCTGGCAGCATTTGCTGTGGCTGCTCTACTGCGTTAATCCAGGCATTTCTGSCNAGGA 720
Qy 241 AlaProAspGluLeu 245
Db 721 GCGCCCGATGAGCTA 735

RESULT 2

US-09-764-877-739
; Sequence 739, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 739
; LENGTH: 1589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-739

Alignment Scores:
Pred. No.: 9.17e-39 Length: 1589
Score: 426.00 Matches: 84
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.1% Indels: 0
DB: 3 Gaps: 0

US-10-073-293A-4 (1-245) x US-09-764-877-739 (1-1589)

Qy 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db 1338 ATGGAAGCCCTACTCCACAGCCTGCTCTCGTTCGGCGACCTTCATGGAAGGATGCAAA 1397
Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40
Db 1398 GACAGTTTACCGATTGTATTAGTTATATTCCGGTGGCCTTTGGCTTCGGTCTGAATGCG 1457
Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyrAlaGly 60
Db 1458 ACCCGTCTGGGATTTCTCTCTCGAAAGCGTTTTTCTCTCGATCATTTATGCGAGGC 1517
Qy 61 AlaSerGlnPheValIleThrAlaMetIleuAlaAlaGlySerSerLeuTrpIleAlaAla 80
Db 1518 GCGAGCCAGTTCGTCAATACCGCATGCTGGCAGCCGGAGTAGTTTGTGGATTGCTGCA 1577
Qy 81 LeuThrValMet 84
Db 1578 CTGACCGTCATG 1589

RESULT 3

US-09-860-670-69
; Sequence 69, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-69

Alignment Scores:
Pred. No.: 9.17e-39 Length: 1589
Score: 426.00 Matches: 84
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.1% Indels: 0
DB: 3 Gaps: 0

US-10-073-293A-4 (1-245) x US-09-860-670-69 (1-1589)

QY 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db 1338 ATGGAAGCCCTACTCCACAGCTGCTCTGTTCCGGGACCTTCATGGAAGGATGCAA 1397
QY 21 AspSerLeuProIleValIleSerTyrlleProValAlaPheAlaPheGlyLeuAsnAla 40
Db 1398 GACAGTTTACCAATTTATTAGTTATATTCGGTGGCCCTTGGCGTTCCGGTCTGAATGCG 1457
QY 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyAlaGly 60
Db 1458 ACCCGTCTGGGATCTCTCTCTCGAAGCGTTTTTTTCTCTGCAATCATTTATGCGAGC 1517
QY 61 AlaSerGlnPheValIleThrAlaMetLeuAlaGlySerSerLeuTrpIleAlaAla 80
Db 1518 GCGAGCCAGTTCTGTCATTCACCGCATGCTGGCAGCCGGAGTAGTTTGTGGATTGCTGCA 1577
QY 81 LeuThrValMet 84
Db 1578 CTGACCGTCATG 1589

RESULT 4

US-09-764-875-55
; Sequence 55, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Number application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 1589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-875-55

Alignment Scores:

Pred. No.: 9.17e-39 Length: 1589
Score: 426.00 Matches: 84
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.1% Indels: 0
DB: 3 Gaps: 0

US-10-073-293A-4 (1-245) x US-09-764-875-55 (1-1589)

QY 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db 1338 ATGGAAGCCCTACTCCACAGCTGCTCTGTTCCGGGACCTTCATGGAAGGATGCAA 1397
QY 21 AspSerLeuProIleValIleSerTyrlleProValAlaPheAlaPheGlyLeuAsnAla 40
Db 1398 GACAGTTTACCAATTTATTAGTTATATTCGGTGGCCCTTGGCGTTCCGGTCTGAATGCG 1457
QY 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyAlaGly 60
Db 1458 ACCCGTCTGGGATCTCTCTCTCGAAGCGTTTTTTTCTCTGCAATCATTTATGCGAGC 1517
QY 61 AlaSerGlnPheValIleThrAlaMetLeuAlaGlySerSerLeuTrpIleAlaAla 80
Db 1518 GCGAGCCAGTTCTGTCATTCACCGCATGCTGGCAGCCGGAGTAGTTTGTGGATTGCTGCA 1577
QY 81 LeuThrValMet 84
Db 1578 CTGACCGTCATG 1589

RESULT 5

US-10-227-646-69
; Sequence 69, Application US/10227646
; Publication No. US20030235829A1

; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4127P1
; CURRENT APPLICATION NUMBER: US/10/227,646
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: US/09/860,670
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-227-646-69

Alignment Scores:

Pred. No.: 9.17e-39 Length: 1589
Score: 426.00 Matches: 84
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.1% Indels: 0
DB: 6 Gaps: 0

US-10-073-293A-4 (1-245) x US-10-227-646-69 (1-1589)

QY 1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db 1338 ATGGAAGCCCTACTCCACAGCTGCTCTGTTCCGGGACCTTCATGGAAGGATGCAA 1397
QY 21 AspSerLeuProIleValIleSerTyrlleProValAlaPheAlaPheGlyLeuAsnAla 40
Db 1398 GACAGTTTACCAATTTATTAGTTATATTCGGTGGCCCTTGGCGTTCCGGTCTGAATGCG 1457
QY 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyAlaGly 60
Db 1458 ACCCGTCTGGGATCTCTCTCTCGAAGCGTTTTTTTCTCTGCAATCATTTATGCGAGC 1517
QY 61 AlaSerGlnPheValIleThrAlaMetLeuAlaGlySerSerLeuTrpIleAlaAla 80
Db 1518 GCGAGCCAGTTCTGTCATTCACCGCATGCTGGCAGCCGGAGTAGTTTGTGGATTGCTGCA 1577
QY 81 LeuThrValMet 84
Db 1578 CTGACCGTCATG 1589

RESULT 6

US-10-242-515-739
; Sequence 739, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496

```
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 739
; LENGTH: 1589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-739

Alignment Scores:
Pred. No.:          9,176-39      Length:      1589
Score:              426.00        Matches:      84
Percent Similarity: 100.0%        Conservative: 0
Best Local Similarity: 100.0%      Mismatches:  0
Query Match:        34.1%         Indels:       0
DB:                 6             Gaps:         0

US-10-073-293A-4 (1-245) x US-10-242-515-739 (1-1589)

Qy      1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db      1338 ATGGAAGCCCTACTCCACAGCCTCTCTGGTTCGGCGACCTTCATGGAAGGATGCAAA 1397

Qy      21 AspSerLeuProIleValIleSerTyriLeProValAlaPheAlaPheGlyLeuAsnAla 40
Db      1398 GACAGTTACCGATTGTTATTAGTTATATCCGGTGGCCTTTGGCTTCGATCG 1457

Qy      41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyralaGly 60
Db      1458 ACCCGTCTGGGATTCTCTCTCTCGAAAGCGTTTTTTTCTCTGCATCATTTATGCAGGC 1517

Qy      61 AlaSerGlnPheValIleThrAlaMetLeuAlaGlySerSerLeuTrpIleAlaAla 80
Db      1518 GCGAGCCAGTTGTCATTACCGCGATGCTGGCAGCCGGAGTAGTTTGTGGATTGCTGCA 1577

Qy      81 LeuThrValMet 84
Db      1578 CTGACCGTCATG 1589

RESULT 7
US-09-764-877-1016/c
; Sequence 1016, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1016
; LENGTH: 1595
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (433)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (467)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-372

Alignment Scores:
Pred. No.:          9,216-39      Length:      1595
Score:              426.00        Matches:      84
Percent Similarity: 100.0%        Conservative: 0
Best Local Similarity: 100.0%      Mismatches:  0
Query Match:        34.1%         Indels:       0
DB:                 3             Gaps:         0

US-10-073-293A-4 (1-245) x US-09-764-875-372 (1-1595)

Qy      1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db      254 ATGGAAGCCCTACTCCACAGCCTCTCTGGTTCGGCGACCTTCATGGAAGGATGCAAA 195

Qy      21 AspSerLeuProIleValIleSerTyriLeProValAlaPheAlaPheGlyLeuAsnAla 40
Db      194 GACAGTTACCGATTGTTATTAGTTATATCCGGTGGCCTTTGGCTTCGATCG 135

Qy      41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyralaGly 60
Db      194 GACAGTTACCGATTGTTATTAGTTATATCCGGTGGCCTTTGGCTTCGATCG 135

US-10-073-293A-4 (1-245) x US-09-764-877-1016 (1-1595)

Qy      1 MetGluSerProThrProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20
Db      254 ATGGAAGCCCTACTCCACAGCCTCTCTGGTTCGGCGACCTTCATGGAAGGATGCAAA 195

Qy      21 AspSerLeuProIleValIleSerTyriLeProValAlaPheAlaPheGlyLeuAsnAla 40
Db      194 GACAGTTACCGATTGTTATTAGTTATATCCGGTGGCCTTTGGCTTCGATCG 135

Qy      41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleTyralaGly 60
Db      194 GACAGTTACCGATTGTTATTAGTTATATCCGGTGGCCTTTGGCTTCGATCG 135

Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%    Mismatches:  0
Query Match:        34.1%        Indels:       0
DB:                 3            Gaps:         0
```

Db 134 ACCGCTCGGAGTCTCTCTCTCGAAGCGTTTTTTTCTCCTCGATCATTTATGCGAGC 75
Qy 61 AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla 80
Db 74 GCGAGCCAGTTCTGTCATTTACCGCATCTGCGAGCCGGAGTAGTTTGTGGATTGCTGCA 15

Qy 81 LeuThrValMet 84
Db 14 CTGACCGTCATG 3

RESULT 9

US-10-242-515-1016/c
; Sequence 1016, Application US/10242515
; Publication No. US20040009488A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005C1

; CURRENT APPLICATION NUMBER: US/10/242,515

; CURRENT FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: 09/764,877

; PRIOR FILING DATE: 2001-01-17

; PRIOR APPLICATION NUMBER: 60/179,065

; PRIOR FILING DATE: 2000-01-31

; PRIOR APPLICATION NUMBER: 60/180,628

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: 60/214,886

; PRIOR FILING DATE: 2000-06-28

; PRIOR APPLICATION NUMBER: 60/217,487

; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/225,758

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/220,963

; PRIOR FILING DATE: 2000-07-26

; PRIOR APPLICATION NUMBER: 60/217,496

; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/225,447

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/218,290

; PRIOR FILING DATE: 2000-07-14

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 4031

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1016

; LENGTH: 1595

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (433)

; OTHER INFORMATION: n equals a,t,g, or c

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (467)

; OTHER INFORMATION: n equals a,t,g, or c

US-10-242-515-1016

Alignment Scores:

Pred. No.: 9.21e-39 Length: 1595

Score: 426.00 Matches: 84

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 34.1% Indels: 0

DB: 6 Gaps: 0

US-10-073-293A-4 (1-245) x US-10-242-515-1016 (1-1595)

Qy 1 MetGluSerProThrGlnProAlaProGlySerAlaThrPheMetGluGlyCysLys 20

Db 254 ATGGAAGCCCTACTCCACGCCCTGCTGTTGGCGACCTTCATGGAGAGTAGCA 195

Qy 21 AspSerLeuProIleValIleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla 40

Db 194 GACAGTTTACCGATTGTTATTATTATATTCGCGTGGCCTTTGGGTTCGGTCTGAATGCG 135
Qy 41 ThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIleIleTyrAlaGly 60
Db 134 ACCGCTCGGAGTCTCTCTCTCGAAGCGTTTTTTTCTCCTGCAATCATTTATGCGAGC 75
Qy 61 AlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrpIleAlaAla 80
Db 74 GCGAGCCAGTTCTGTCATTTACCGCATCTGCGAGCCGGAGTAGTTTGTGGATTGCTGCA 15
Qy 81 LeuThrValMet 84
Db 14 CTGACCGTCATG 3

RESULT 10

US-09-070-927A-242/c

; Sequence 242, Application US/09070927A

; Patent No. US20020120116A1

; GENERAL INFORMATION:

; APPLICANT: Charles A. Kunsch

; PATENT ATTORNEY: Steven Barash

; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides

; NUMBER OF SEQUENCES: 982

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/070,927A

; FILING DATE: 04-May-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/046,655

; FILING DATE: 1997-05-16

; APPLICATION NUMBER: 60/044,031

; FILING DATE: 1997-05-06

; APPLICATION NUMBER: 60/066,009

; FILING DATE: 1997-11-14

; ATTORNEY/AGENT INFORMATION:

; NAME: Kenley K. Hoover

; REGISTRATION NUMBER: 40,302

; REFERENCE/DOCKET NUMBER: PB369

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 242:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12445 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 242:

US-09-070-927A-242

Alignment Scores:

Pred. No.: 9.19e-16 Length: 12445

Score: 238.50 Matches: 71

Percent Similarity: 46.4% Conservative: 33

Best Local Similarity: 31.7% Mismatches: 99

Query Match: 19.1% Indels: 21

DB: 3 Gaps: 7

US-10-073-293A-4 (1-245) x US-09-070-927A-242 (1-12445)

Qy 18 GlyCysLysAspSerLeuProileValIleSerTyrlleProValAlaPheAlaPheGly 37
 Db 5700 GGCATTCGTGAAGCCCTTCCCAAGGTTTGGCTATATGGTATTTGGCATTCGTTTGGT 5641
 Qy 38 LeuAenAlaThrArgLeuGlyPheSerProLeuGluSerValPhePheSerCysIleIle 57
 Db 5640 ATCGTAGGACCAAGCAGCAGGTTTTTCCGCTCTTATTGTTAGTGCATGTCGTTTTTTATT 5581
 Qy 58 TyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGlySerSerLeuTrp 77
 Db 5580 TTTGAGGATCTGCTCAGTTGTAAAGTGTAGTATGTTGACTGGCGGAGCCCTTCTTA 5521
 Qy 78 IleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyrlleGlyProSerLeu 97
 Db 5520 TCTATTGTTAGCGACTTCTTAGTAATGCGGTATGATCTTATGGGATGACCAT 5461
 Qy 98 ArgSerArgIle---IleGlnArgLeuGlnIlySerIlySerIlySerIlySerIlySerIly 116
 Db 5460 GCGCCTTACTTTAAAGCCGAATCTCTCGGTAAAT-----CTCTGG---TTGGGA 5413
 Qy 117 -----LeuThrAspGluValPheAlaAlaAlaThrAlaAlaLeuValArgAsnArg 134
 Db 5412 ACCTTACTGACGGATGAAGTTTTCGCTAGGAATGAATAATAAACACACCAAGGA 5353
 Qy 135 ArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerTrpValPhe 154
 Db 5352 CGTTTAAGTTTGAAGTGTAAATGCTGCCAATTTGATTTCTATGCGACTGGGTTTT 5293
 Qy 155 GlyThrValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrlleProAlaValGlu 174
 Db 5292 TCGAGCATCATCGGTCTATCTAGGCGC-----TTTATGCAACCCCT 5248
 Qy 175 AlaAlaLeuGly-----PheMetLeuProAlaLeuPheMetSerPheLeuLeuAlaSer 192
 Db 5247 CAAGCATTAGGCTTTGAGTTTGGCGTCTGCTGCTATGTTTCATTTGGTTTACTCTATCTGCAG 5188
 Qy 193 PheGlnArgLysGlnSerLeuCysValThrAlaAlaLeuValGlyAlaLeuAlaGlyVal 212
 Db 5187 ATAATTAGTATCGTTCATGAAATCGTTTACAACTCATCATGCTCTGATTAATTTT 5128
 Qy 213 ThrLeuPheSerIle-----ProValAlaIleLeuAlaGlyIle 225
 Db 5127 GGTTTGATGCTATTTGGTCTAATTTTATCCAGTAATTTAAATGTCYTAGTTGKACG 5068
 Qy 226 ValCysGlyCys 229
 Db 5067 TTAATCGGGTGT 5056
 RESULT 11
 US-10-724-972A-2949
 ; Sequence 2949, Application US/10724972A
 ; Publication No. US20040147734A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Doucette-Stamm, Lynn
 ; APPLICANT: Bush, David
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: PATH03-16
 ; CURRENT APPLICATION NUMBER: US/10/724, 972A
 ; PRIOR FILING DATE: 2003-12-01
 ; PRIOR APPLICATION NUMBER: 09/450, 969
 ; PRIOR FILING DATE: 1999-11-29
 ; PRIOR APPLICATION NUMBER: 09/134, 001
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: 60/064, 964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: 60/055, 779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 7544
 ; SEQ ID NO 2949
 ; LENGTH: 729
 ; TYPE: DNA
 ; ORGANISM: S.epidermidis

US-10-724-972A-2949
 Alignment Scores:
 Pred. No.: 5.92e-16 Length: 729
 Score: 225.50 Matches: 66
 Percent Similarity: 45.5% Conservative: 39
 Best Local Similarity: 28.6% Mismatches: 89
 Query Match: 18.1% Indels: 37
 DB: 7 Gaps: 7
 US-10-073-293A-4 (1-245) x US-10-724-972A-2949 (1-729)
 Qy 14 ThrPheMetGluGlyCysLysAspSerLeuProileValIleSerTyrlleProValAla 33
 Db 49 ACGTTTAAACAAGGTGTGAAGAGTGTATCCCACGTTACTTGGTTATGCTGGGTAGGA 108
 Qy 34 PheAlaPheGlyLeuAenAlaThrArgLeuGlyPheSerProLeuGluSerValPhePhe 53
 Db 109 CTATCGTTTGAATTTGGCAGCCCTCCCAAAATTTTCAGTGTGTTTTAGAAAATTTATTTATTG 168
 Qy 54 SerCysIleIleTyrlleGlyAlaSerGlnPheValIleThrAlaMetLeuAlaAlaGly 73
 Db 169 TGTCTGATTTATTTATGCTGGTGCAGCCCAATTTATTTATTTGTACATGATGATGTCAGGC 228
 Qy 74 SerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyrl 93
 Db 229 ACCCTATTTCGCAATTTGCTTACACACTTATCGTTAACTCTCGAATGTTCTTATTA 288
 Qy 94 GlyProSerLeuArgSerArgIleIleGlnArgLeuGlnIlySerIlySerIlySerIly 113
 Db 289 AGTATGACTTTAGCACCAATTTATAAGCAATATGATTTTGGATAGGTAGGCGTTTGA 348
 Qy 114 AlaPheGlyLeuThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnArg 133
 Db 349 ACGTTA---TTAACAGATGAAACTTTTGGAGTTGCTATAACGCCATATGTCAAA---GGT 402
 Qy 134 ArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerSerTrpVal 153
 Db 403 GAAAAAATTAAAGATCGATCGCTACAGGATAAATATTACTGCTTACTTATTTTGGACT 462
 Qy 154 PheGlyThrValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrlleProAlaVal 173
 Db 463 GTTTTCATGTGTAATCGGTGCCATTTTCGGAGAG-----TATATTTCAAAAT 507
 Qy 174 GluAlaAlaLeuGly-----PheMetLeuProAlaLeuPheMetSerPheLeuLeuAla 191
 Db 508 CCTGATCGCTTGGCCTAGACTTTTGCCATTACCGCAATGTTTATTTTATGATATATCT 567
 Qy 192 SerPheGln-----ArgLysGlnSerLeu-----CysVal 201
 Db 568 CAATTTGAGGGATTAGAAATCAGATTGAGATATATATTGTACTCATGTTATGTTGTG 627
 Qy 202 -----ThrAla 203
 Db 628 ATGTGATGATGCTTCTTAAAGTTCAATTTACCTTCACTCATACGTAGCAAAATTTAATAGCC 687
 Qy 204 AlaLeuValGlyAlaLeuAlaGlyValThrLeu 214
 Db 688 GCAATTTGCTGCTGCAATTTTAGGGGCTGTTGATG 720
 RESULT 12
 US-10-398-221-8/c
 ; Sequence 8, Application US/10398221
 ; Publication No. US20040018514A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KUNST, Frederik
 ; APPLICANT: GLASER, Philippe
 ; TITLE OF INVENTION: Listeria innocua, genome and applications
 ; FILE REFERENCE: 344 702 - US
 ; CURRENT APPLICATION NUMBER: US/10/398, 221
 ; PRIOR FILING DATE: 2003-03-27
 ; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
 ; PRIOR FILING DATE: 2001-10-04


```
Db 1483576 ATTTTATTAAATCGTGTATTATGCGTTTATGACACCAAGATTAGCAATTTATGCGCTACC 1483635
Qy 226 ValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGln 239
Db 1483636 TTACTTGTTGTTTGTGATTGGAGTGATAGTTGAAAAATGGCAA 1483677

RESULT 14
US-10-398-221-3752
; Sequence 3752, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederick
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3752
; LENGTH: 6157
; TYPE: DNA
; ORGANISM: Listeria monocytogenes 4b
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
US-10-398-221-3752

Alignment Scores:
Pred. No.: 2,19e-14 Length: 6157
Score: 223.00 Matches: 70
Percent Similarity: 44.6% Conservative: 38
Best Local Similarity: 28.9% Mismatches: 110
Query Match: 17.9% Indels: 24
DB: 7 Gaps: 8

US-10-073-293A-4 (1-245) x US-10-398-221-3752 (1-6157)
Qy 14 ThrPheMetGluGlyCysLeuAspSerLeuProIleValIleSerTyrlleProValAla 33
Db 4269 AGTTTTATGATGGTGTAGAGCGCTCTCTACCGTCTTGGTGGTATGCTGGGATGGT 4328
Qy 34 PheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePhe 53
Db 4329 ATCGCTGCAGGATAGTAGGAAAGCATCCCAATTTAAGCCCTTTTGAAGTGACGCTACTT 4388
Qy 54 SerCysIleIleTyAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaGly 73
Db 4389 CGGATTATCGTTTATGACGGTGGCGGCAATTTATTTCTGTTGTTTATTACTACAA 4448
Qy 74 SerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTy 93
Db 4449 AGTCGATATCAGCAATCATTTTACTACTTTTAAATTAATTAATCAAGCAATTTTAAATG 4508
Qy 94 GlyProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerIleThrAlaLeuTrp 113
Db 4509 AGT-----ATGCGGAAGCTCCTCAATTTTAAAGAAATTTCTTTGTGG 4550
Qy 114 -----AlaPheGly-----LeuThrAspGluValPheAlaAlaAlaThrAlaLys 128
Db 4551 AATAATATTGGATTGGTGGCCCTTCTGACGATGAACGTTTGGGGTTCGATGATCAG 4610
Qy 129 LeuValArgAsnAsnArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSer 148
Db 4611 ---ATTGGTAATAAAAAACCTGTTAGTGTAAATGGATGCACGGAATAAATGTAACCTGCC 4667
Qy 149 TrpSerSerTrpValPheGlyThrValIleGlyAlaPheSerGlySerGlyLeuLeuGln 168
```

```
Db 4668 TACTTGGCATGATTGTGCTTGTATCGTTGGGCAATTTATTGGCAACTGGCTC----- 4721
Qy 169 GlyTyTrProAlaValGluAla---AlaLeuGlyPheMetLeuProAlaLeuPheMetSer 187
Db 4722 -----CCAAATCCAGAACAAATTTGGTTTATAGACTTTCGCTGTGCTGCAATGTTTATCGGT 4775
Qy 188 PheLeu---LeuAlaSerPheGlnArgLysGlnSerLeuCysValThrAlaAlaLeuVal 206
Db 4776 TTGCTATATTATCAAGTAGTAGTGATATAAAGCAAAAAAATAAGTAGAGTTTGTTCGTC 4835
Qy 207 GlyAlaLeuAlaGlyValThrLeu-----PheSerIleProValAla 220
Db 4836 ATGATACTAGTCGCTGATTTTAAATCTCTTTATGCTGTGTGATGACGCCAGAACTAGCT 4895
Qy 221 IleLeuAlaGlyIleValCysGlyCysLeuThrAlaLeuIleGlnAlaPheTrpGlnGly 240
Db 4896 ATTTTAACCTGCACCTTACTCGGATGTTTGATGGGAGTGATTATAGAGATGCGGTAA 4955
Qy 241 AlaPro 242
Db 4956 GTTCCT 4961

RESULT 15
US-08-781-986A-136/c
; Sequence 136, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 136:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11823 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-136

Alignment Scores:
Pred. No.: 7,58e-12 Length: 11823
Score: 204.50 Matches: 65
Percent Similarity: 47.0% Conservative: 45
Best Local Similarity: 27.8% Mismatches: 97
Query Match: 16.4% Indels: 27
DB: 2 Gaps: 10
```


Search completed: February 16, 2006, 18:23:42
Job time : 1786.97 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:47:13 ; Search time 655.857 Seconds
(without alignments)
792.964 Million cell updates/sec

Title: US-10-073-293A-4
Perfect score: 1248
Sequence: 1 MESPTQPAPGSAFMFGCK.....VCGCLTALIOAQWGAPDEL 245

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7204252 seqs, 1061369211 residues

Total number of hits satisfying chosen parameters: 14408504

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ASSWEB.spool/US10073293/runat.15022006.120743.1175/app.query.fasta.1
-DB=Published Applications NA New -QFMT=fastp -SURFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pt0 -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs03p
-USER=US10073293 -CGN 1.1.579 -runat.15022006.120743.1175 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMBOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

- 1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	229.5	18.4	693	8	US-10-793-626-267
2	229.5	18.4	3058	8	US-10-793-626-3585
3	229.5	18.4	3077	8	US-10-793-626-3594
4	108	8.7	4381	12	US-11-136-527-2390

5	105	8.4	1035	11	US-11-082-389-13	Sequence 13, Appl
c 6	93	7.5	2748	8	US-10-467-657-5231	Sequence 5231, Ap
c 7	93	7.5	2763	8	US-10-467-657-7353	Sequence 7353, Ap
8	91.5	7.3	1299	8	US-10-873-528-229	Sequence 229, App
9	91.5	7.3	1947	8	US-10-793-626-1059	Sequence 1059, Ap
c 10	91.5	7.3	2397	8	US-10-793-626-3342	Sequence 3342, Ap
c 11	91.5	7.3	3298	8	US-10-793-626-3792	Sequence 3792, Ap
c 12	91.5	7.3	3332	12	US-11-136-527-74	Sequence 74, Appl
c 13	90.5	7.3	1419	8	US-10-467-657-1873	Sequence 1873, Ap
14	90	7.2	110711	8	US-10-995-561-13264	Sequence 13264, A
15	89.5	7.2	168516	12	US-11-121-086-3	Sequence 3, Appli
16	87.5	7.0	1706	12	US-11-012-668-3	Sequence 3, Appli
c 17	87.5	7.0	3639	8	US-10-793-626-3393	Sequence 3393, Ap
18	87	7.0	1416	8	US-10-467-657-2267	Sequence 2267, Ap
19	87	7.0	168753	12	US-11-181-234-1	Sequence 1, Appli
c 20	86.5	6.9	4041	8	US-10-793-626-3573	Sequence 3573, Ap
c 21	86.5	6.9	67858	8	US-10-995-561-13484	Sequence 13484, A
c 22	86.5	6.9	124972	12	US-11-121-086-100	Sequence 100, App
c 23	86	6.9	903	12	US-11-098-686-9238	Sequence 9238, Ap
c 24	86	6.9	2469	8	US-10-467-657-2525	Sequence 2525, Ap
c 25	86	6.9	33931	12	US-11-124-367A-5006	Sequence 5006, Ap
c 26	86	6.9	1457619	12	US-11-098-686-8739	Sequence 8739, Ap
c 27	85.5	6.9	1362	8	US-10-793-626-3259	Sequence 3259, Ap
c 28	85.5	6.9	3181	8	US-10-793-626-3673	Sequence 3673, Ap
c 29	85.5	6.9	88421	12	US-11-205-109-1	Sequence 1, Appli
c 30	85	6.8	1072	6	US-09-925-065A-87567	Sequence 87567, A
c 31	84.5	6.8	2089	12	US-11-210-316-25	Sequence 25, Appl
c 32	84	6.7	2214	11	US-11-082-389-361	Sequence 361, App
c 33	83.5	6.7	1524	8	US-10-467-657-7825	Sequence 7825, Ap
c 34	83.5	6.7	3158	8	US-10-821-234-175	Sequence 175, App
c 35	83	6.7	697	6	US-09-925-065A-876781	Sequence 876781, A
36	83	6.7	1452	8	US-10-467-657-3819	Sequence 3819, Ap
37	83	6.7	1617	8	US-10-467-657-3145	Sequence 3145, Ap
38	83	6.7	1910	12	US-11-136-527-1918	Sequence 1918, Ap
c 39	82.5	6.6	1425	12	US-11-074-176-335	Sequence 335, App
c 40	82.5	6.6	1464	12	US-11-074-176-125	Sequence 125, App
c 41	82.5	6.6	14896	9	US-11-245-147-243	Sequence 243, App
c 42	82.5	6.6	14896	12	US-11-000-688-946	Sequence 946, App
c 43	82.5	6.6	14896	12	US-11-076-427A-31	Sequence 31, Appl
c 44	82	6.6	1434	6	US-09-925-065A-705332	Sequence 705332, A
c 45	82	6.6	2079	8	US-10-467-657-6175	Sequence 6175, Ap

ALIGNMENTS

RESULT 1
US-10-793-626-267
; Sequence 267, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; PRIOR FILING DATE: 2004-03-04
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 267
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-267

Alignment Scores:
Pred. No.: 5.63e-14 Length: 693
Score: 229.5 Matches: 65
Percent Similarity: 49.3% Conservative: 47
Best Local Similarity: 28.6% Mismatches: 102

Query Match: 18.4% Indels: 13
DB: 8 Gaps: 6

US-10-073-293A-4 (1-245) x US-10-793-626-267 (1-693)

Qy 14 ThrPheMetGluGlyCysLysAspSerLeuProIleValIleSerTyrIleProValAla 33
Db 13 ACGTTTAAACAAGGTGTAAGAGTGTATCCACGTTACTTGGTTATGACAGGTAGGA 72

Qy 34 PheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePhe 53
Db 73 CTATCGTTGGAATTTGGCAGTCTCCCAAAATTTCAAGTGTATTTAGAAATTTATTATTG 132

Qy 54 SerCysIleIleTyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaGly 73
Db 133 TGTCTGATTTATGCTGGTGGCAGCTCAATTTATTTTGTACATTAGTGCAGGC 192

Qy 74 SerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyr 93
Db 193 ACCCTATTTCTGCAATTTGTCTACATATCTTATCGTTAACTCTCGAATGTTCTTATTA 252

Qy 94 GlyProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrp 113
Db 253 AGTATGACTTTAGCACCTAAATATAAGCAATATGGAATTTTGGAAATAGGTAGGCTTGA 312

Qy 114 AlaPheGlyLeuThrAspGluValPheAlaAlaThrAlaLysLeuValArgAsnAsn 133
Db 313 ACGTTA---TTAACAGATGAAACTTTTGGCGTTCGTATACACCATATGTTTAA---GGT 366

Qy 134 ArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerTrpVal 153
Db 367 GAAAAAATTACGATCGATGCTACACGCACTAATATTTACTGCTTACTTATTTGGACT 426

Qy 154 PheGlyThrValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrProAlaVal 173
Db 427 GTTTCCTGTGTAATCGGTGCAATTTTCGGAGAG-----TATATTTCAAAT 471

Qy 174 GluAlaAlaLeuGly-----PheMetLeuProAlaLeuPheMetSerPheLeuLeuAla 191
Db 472 CCGTATGCGCTCGCGCTAGACTTTTGCAATTACCGCAATGTTTATTTTATGATATCT 531

Qy 192 SerPheGln-----ArgLysGlnSerLeuCysValThrAlaAlaLeuValGlyAlaLeu 209
Db 532 CAATTTGAAGGATTAAGAAATCAGATTGAGATATATATTGTACTCATTTGTATGTGTG 591

Qy 210 AlaGlyValThrLeuPhe-----SerIleProValAlaIleLeuAlaGlyIleValCys 227
Db 592 ATTGTGATGATGCTTCTTCTAAGTTCATTTCTACCTTACCTAGCAATTTTAAATAGCC 651

Qy 228 GlyCysLeuThrAlaLeuIle 234
Db 652 GCAATTTGTTGCTGCAATTTGTTA 672

RESULT 2

US-10-793-626-3585/c
; Sequence 3585, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3585
; LENGTH: 3058
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic

i OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3585

Alignment Scores:
Pred. NO.: 3.18e-13 Length: 3058
Score: 229.50 Matches: 65
Percent Similarity: 49.3% Conservative: 47
Best Local Similarity: 28.6% Mismatches: 102
Query Match: 18.4% Indels: 13
DB: 8 Gaps: 6

US-10-073-293A-4 (1-245) x US-10-793-626-3585 (1-3058)

Qy 14 ThrPheMetGluGlyCysLysAspSerLeuProIleValIleSerTyrIleProValAla 33
Db 1397 ACGTTTAAACAAGGTGTAAGAGTGTATTTCCACGTTACTTGGTTATGACAGGTAGGA 1338

Qy 34 PheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePhe 53
Db 1337 CTATCGTTGGAATTTGGCAGTCTCCCAAAATTTCAAGTGTATTTAGAAATTTATTATTG 1278

Qy 54 SerCysIleIleTyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaGly 73
Db 1277 TGTCTGATTTATGCTGGTGGCAGCTCAATTTATTTTGTACATTAGTGCAGGC 1218

Qy 74 SerSerLeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyr 93
Db 1217 ACCCTATTTCTGCAATTTGTCTTACATATCTTATCGTTAACTCTCGAATGTTCTTATTA 1158

Qy 94 GlyProSerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrp 113
Db 1157 AGTATGACTTTAGCACCTAATATAAGCAATATGGAATTTTGGAAATAGGTAGGCTTGA 1098

Qy 114 AlaPheGlyLeuThrAspGluValPheAlaAlaThrAlaLysLeuValArgAsnAsn 133
Db 1097 ACGTTA---TTAACAGATGAACTTTTGGCGTTCGTATACACCATATGTTTAA---GGT 1044

Qy 134 ArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrpSerTrpVal 153
Db 1043 GAAAAAATTACGATCGATGCTACACGCACTAATATTTACTGCTTACTTATTTGGACT 984

Qy 154 PheGlyThrValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrProAlaVal 173
Db 983 GTTTCCTGTGTAATCGGTGCAATTTTCGGAGAG-----TATATTTCAAAT 939

Qy 174 GluAlaAlaLeuGly-----PheMetLeuProAlaLeuPheMetSerPheLeuLeuAla 191
Db 938 CCGTATGCGCTCGCGCTAGACTTTTGCAATTACCGCAATGTTTATTTTATGATATCT 879

Qy 192 SerPheGln-----ArgLysGlnSerLeuCysValThrAlaAlaLeuValGlyAlaLeu 209
Db 878 CAATTTGAAGGATTAAGAAATCAGATTGAGATATATATTGTACTCATTTGTATGTGTG 819

Qy 210 AlaGlyValThrLeuPhe-----SerIleProValAlaIleLeuAlaGlyIleValCys 227
Db 818 ATTGTGATGATGCTTCTTCTAAGTTCATTTCTACCTTACCTAGCAATTTTAAATAGCC 759

Qy 228 GlyCysLeuThrAlaLeuIle 234
Db 758 GCAATTTGTTGCTGCAATTTGTTA 738

RESULT 3
US-10-793-626-3594
; Sequence 3594, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09

Qy 42 ArgLeuGly-----PheSerProLeuGluSerValPhe-----Phe 53
 Db 2115 CGGTTCCGGCGCGGTTTTCGGCTGCTGAACGCTGTTTCGGCAACGGCGCGGTTTC 2056
 Qy 54 SerCysIleIleTyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaGly 73
 Db 2055 CGGTGC-----GGTTTGGCCACGTTTCGGCTGCGCTGCGCGCGGTTT 2008
 Qy 74 SerSerLeuTrpIle-----AlaAlaLeuThr 82
 Db 2007 TCCGGTTTGACCTTCGGTTTCGGCGGTGCGCATCTGCAACGGGTGCGGAGCTGTAC 1948
 Qy 83 ValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArgIleIle 102
 Db 1947 GTTGGCGCTTGGATTTCGGCTTCGTT----- 1921
 Qy 103 GlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGluValPhe 122
 Db 1920 -----GGCGCGTTTCGGTGGCACGGTCGCGCGGTTTC 1891
 Qy 123 AlaAlaAlaThrAlaLysLeuValArgAsn----- 132
 Db 1890 GTTCCGGCGCGGTTTCGGTTTCGGCTTCGGCTTCGGCGCGGCTTTCGTCTGC 1831
 Qy 133 -----AsnArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrp 149
 Db 1830 GGCAGTTTTCGCCCGCTTCGGCGACTTCTATTTCGTCTTCGGCTTCGCTTTCGCT 1777
 Qy 150 SerSerTrpValPheGlyThrValIleGlyAlaPheSer----- 162
 Db 1776 GCGCGCGCGGTTTTCGGCGCGGTTTTCGGCGCGCTTCGGCTTCGGCTTCGCTTCG 1717
 Qy 163 -----GlySerGlyLeuGlnGlyTyrProAlaValGluAlaAlaLeu 177
 Db 1716 TTTTTCGAGGTTTCGGCAGCGGTACGCTTGGCTTCGGCTTCGGCTTCGCTTCGCTTC 1657
 Qy 178 -----GlyPheMetLeu 181
 Db 1656 GAGCCATGCTTTGAAGCTGTCCTCCACGAGGAGTTTTCCTC 1615

RESULT 7

US-10-467-657-7353/c
 ; Sequence 7353, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SpA
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MASIGNANI Vega
 ; APPLICANT: MONACI Elisabetta
 ; TITLE OF INVENTION: CONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/467,657
 ; PRIOR FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: GB-0103424.8
 ; PRIOR FILING DATE: 2001-02-12
 ; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: SeqWin99, version 1.04
 ; SEQ ID NO 7353
 ; LENGTH: 2763
 ; TYPE: DNA
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-7353

Alignment Scores:
 Pred. No.: 29.2 Length: 2763
 Score: 93.00 Matches: 54
 Percent Similarity: 33.2% Conservative: 17
 Best Local Similarity: 25.2% Mismatches: 57
 Query Match: 7.5% Indels: 86
 DB: 8 Gaps: 10

US-10-073-293A-4 (1-245) x US-10-467-657-7353 (1-2763)

Qy 27 IleSerTyrIleProValAlaPheAlaPheGlyLeuAsnAla-----Thr 41
 Db 2190 GTCGGTTCCGGCGGTTTCGGATTTCGGATTTCGGCTTCGCTTCGCGGTATA 2131
 Qy 42 ArgLeuGly-----PheSerProLeuGluSerValPhe-----Phe 53
 Db 2130 CGGTTCCGGCGCGGTTTTCGGCTGCTGAACGGCTTTCGGCAACGGCGCGGTTTC 2071
 Qy 54 SerCysIleIleTyrAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAlaGly 73
 Db 2070 CGGTGC-----GGTTTGGCCACGTTTCGGCTTCGGCTTCGGCTTCGCGCGGTTT 2023
 Qy 74 SerSerLeuTrpIle-----AlaAlaLeuThr 82
 Db 2022 TCCGGTTTGACCTTCGGTTTCGGCGGTGCGCATCTGCAACGGGTGCGGAGCTGTAC 1963
 Qy 83 ValMetAlaMetAspValArgHisValLeuTyrGlyProSerLeuArgSerArgIleIle 102
 Db 1962 GTTCCGGCTTGGATTTCGGCTTCGTT----- 1936
 Qy 103 GlnArgLeuGlnLysSerLysThrAlaLeuTrpAlaPheGlyLeuThrAspGluValPhe 122
 Db 1935 -----GGCGCGTTTCGGTGGCACGGTCGCGCGGTTTC 1906
 Qy 123 AlaAlaAlaThrAlaLysLeuValArgAsn----- 132
 Db 1905 GTTCCGGCGCGGTTTCGGTTTCGGCTTCGGCTTCGGCTTCGGCTTCGCTTCGTC 1846
 Qy 133 -----AsnArgArgTrpSerGluAsnTrpMetIleGlyIleAlaPheSerSerTrp 149
 Db 1845 GGCAGTTTTCGCCCGCTTCGGCGACTTCTATTTCGTCTTCGGCTTCGCTTCGCTTC 1792
 Qy 150 SerSerTrpValPheGlyThrValIleGlyAlaPheSer----- 162
 Db 1791 GCGCGCGGTTTTCGGCGCGGTTTTCGGCGCGCTTCGGCTTCGGCTTCGCTTCGCTTC 1732
 Qy 163 -----GlySerGlyLeuGlnGlyTyrProAlaValGluAlaAlaLeu 177
 Db 1731 TTTTTCGAGGTTTCGGCAGCGGTACGCTTGGCTTCGGCTTCGGCTTCGCTTCGCTTC 1672
 Qy 178 -----GlyPheMetLeu 181
 Db 1671 GAGCCATGCTTTGAAGCTGTCCTCCACGAGGAGTTTTCCTC 1630

RESULT 8

US-10-873-528-229
 ; Sequence 229, Application US/10873528
 ; Publication No. US20050276814A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Microbial Technics Limited
 ; APPLICANT: Gilbert, Christophe FG
 ; APPLICANT: Hansbro, Philip M
 ; TITLE OF INVENTION: Proteins
 ; FILE REFERENCE: PWC/21129WO
 ; CURRENT APPLICATION NUMBER: US/10/873,528
 ; CURRENT FILING DATE: 2004-06-23
 ; PRIOR APPLICATION NUMBER: US/09/769,787
 ; PRIOR FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: GB 9816337.1
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: US 60/125164
 ; PRIOR FILING DATE: 1999-03-19
 ; NUMBER OF SEQ ID NOS: 388
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 229
 ; LENGTH: 1299
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 US-10-873-528-229

Alignment Scores:

Pred. No.: 17.2 Length: 1299


```

Db      1117  -----|||:::||||:::|||||
Qy      213  -----TTATTAGAGGGGTTACTGGTGAATGCTT 1146
Db      1147  GCAATGGCATCAGCTGACTATCAATATACACAACATTATTCTTAGTAGCTCAGTCCAC 1206
Qy      219  ValAlaLeuAlaGlyIleValCysGlyThrAlaLeuIleGlnAlaPheTrp 238
Db      1207  TATACATTGGTTACTGGTGTAGTATTGGCCTGCTTAGCTGTTTAATC-----TTCTGG 1260

RESULT 10
US-10-793-626-3342/c
; Sequence 3342, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3342
; LENGTH: 2397
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3342

Alignment Scores:
Pred. No.:          35.2      Length:      2397
Score:             91.50     Matches:    63
Percent Similarity: 32.3%    Conservative: 34
Best Local Similarity: 21.0%  Mismatches: 98
Query Match:       7.3%     Indels:    105
DB:                8       Gaps:      15

US-10-073-293A-4 (1-245) x US-10-793-626-3342 (1-2397)

Qy      9  AlaProGlySerAlaThrPheMetGluGlyCysLysAspSerLeuProIleValIleSer 28
Db      2066 GCACCATTGCTGGTGAATTCAGTCCAGT-----CCAGGTGTTACTAT 2022
Qy      29  TyrIleProValAlaPheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeu 48
Db      2021 TATTAAATTGCAATTCAAAATATCTCGTATCGGATCGTTAATGACTGGTATCAAC-----1968
Qy      49  GluSerValPhePheSerCysIleIleTyrAla-----59
Db      1967 -----TTCTTGTGTACGATCTTAAGATGTAAAACTCCAAACAATGAAGTTTATGCAA 1917
Qy      60  -----GlyAlaSerGlnPheValIleThrAlaMetLeuAlaGlySerSer 75
Db      1916 ATGCCAATGTCAGTGTACAAACAATCTACACACATTAATCGTTATATTAGCAATGCCCA 1857
Qy      76  LeuTrpIleAlaAlaLeuThrValMetAlaMetAspValArgHisValLeuTyrGlyPro 95
Db      1856 GTGTTCAGCTAGACACTTGTCTTTAATGACTGCTGAT-----AGAAATTTTGGT---1809
Qy      96  SerLeuArgSerArgIleIleGlnArgLeuGlnLysSerLysThrAlaLeuTrp-----113
Db      1808 -----ACTCAGTCTCTTCTCAGTGTAGCAAAATGGCGGTATGCTTGGCAAAAC 1758
Qy      114  -----114
Db      1757 TTCTTCGTGGGTATGGGGGCACCCCTGAAGTTTATATCGTTATTTTGGCAGCAATTCGGTATG 1698
Qy      118  ThrAspGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTrpSer 137

```



```
Db 288 CCCGTCCTCGCGTCTGAATCTCTGTTTTCGCTGACGATGCGCGCGCGGATCGGGCT 347
Qy 25 eVal-----lleSerTyrlleProValAlaPhe-----AlaPheGlyLeuAs 39
Db 348 GGTGCTGTTTGAACGTAGCTCTCTGCGCCATCGCTTCTTACTGTTTTCGTATGGA 407
Qy 39 nAlaThrArgLeuGlyPheSerProLeuGluSerVal----- 51
Db 408 AGGCGCGCCCTTCTGCTTCGCGCGCAACTCGTGCCCAACTCGCCATTCGTCTGCT 467
Qy 52 -----PhePheSerCysllelleTyrlleAlaGlyAlaSerGlnPheValilleThralaMe 69
Db 468 GTTCCGCTGACGTCGGCTGCTGCACTTTCCGCGCGCAACCTCCGCTCCTGACCGCGT 527
Qy 69 tLeuAlaAlaGlySerSerLeuTrpilleAlaAlaLeuThrValMetAlaMetAsp----- 87
Db 528 TTACGCGCTGCGCAAC-----CITGCGCGCGCGCGCTTTTGTGTTTCAAACCG 578
Qy 88 -----ValArgHisValLeuTyrlleGlyProSerLeuArgSerArgilleIl 102
Db 579 ATGCCGCTGTGAAGCGCGCTGCGCGCGCGCGCTTTTTCGCGCGCGCTCTGTCACCGCGGCT 638
Qy 102 eGln-----ArgLeuGlnLysSerLysThrAlaLeuThrAlaPheGlyLeuThrAs 119
Db 639 GCGTACGCGATACCGCTCGGCACGTAGCAGCCTTGCCTATTGGGGGCTGGCATCCGCGGA 698
Qy 119 pGluValPheAlaAlaAlaThrAlaLysLeuValArgAsnAsnArgArgTrpSerGluAs 139
Db 699 CCGTTTGTCTGAAAAAATATGCGGCGCTGGACACGTCCGC-----GT 743
Qy 139 nTrpMetilleGlylleAlaPheSerSer----- 148
Db 744 TTATTGATGCTATTCTGTTGCGCGCGCGCGCATTTATGTCCAAAGCATCTTTTCAAC 803
Qy 149 ----TrpSerSerTrpValPheGlyThrVal----- 157
Db 804 GGTCTGACACCGGTATTTTTCGTCGCAATGAAAGAAACGCGCGCGCGCTCTC 863
Qy 158 -----lleGlyAlaPhe 161
Db 864 GGCAACGCGAGATCCGCGCGCGCTGCTTGCCTCGCGCTCTGCGCTGACCGGAATTTT 923
Qy 161 eSer-----GlySerGlyLeuLeu-----GlnGlyTyrlleProAlaValGlu-----AlaAlaLe 177
Db 924 CTCGCCCTCGCTCTCTCTGCTGCGGAAACATGACGCGCGCTCGGTTTACCGCTGT 983
Qy 177 uGlyPheMetLeuProAlaLeuPhe-----MetSerPheLeuLeuAlaSe 192
Db 984 ATCGTGATGCTGCGCGCGCTGTTTATACGCTGACCGGAATCAGCGCATCGGTTTGA 1043
Qy 192 rPheGlnArgLysGlnSerLeuCysValThrAlaAlaLeuValGlyAlaAlaGlyVa 212
Db 1044 CGTCTCGCGCAA---ACGCGTCGATGCGCGCTGCGCACCTTGGCGCGCTGGCG---GC 1097
Qy 212 lThrLeuPheSerlleProValAlaAlaLeuAlaGlylleValCysGlyCysLeuThral 232
Db 1098 AAACCTGCTGCTGCGCGCTGCGTACCGTCCGCGCGGACGCGCGCGCGCGGCTGC 1157
Qy 232 aLeulleGlnAlaPheTrp 238
Db 1158 CTGTGCGCGCTCATCTGG 1176
```

RESULT 14

```
US-10-995-561-13264
; Sequence 13264, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
```

```
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13264
; LENGTH: 110711
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(110711)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-995-561-13264
```

Alignment Scores:

```
Pred. No.: 4.4e+03 Length: 110711
Score: 90.00 Matches: 61
Percent Similarity: 39.1% Conservative: 36
Best Local Similarity: 24.6% Mismatches: 104
Query Match: 7.2% Indels: 48
DB: 8 Gaps: 12
```

US-10-073-293A-4 (1-245) x US-10-995-561-13264 (1-110711)

```
Qy 14 ThrPheMetGluGlyCysLysAspSerLeuProLleValleSerTyrlleProValAla 33
Db 24165 ACTCTCTCTCATCTTGT-----CATCTCATTATT 24197
Qy 34 PheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerValPhePhe 53
Db 24198 CATTACTTTGGTCTTCAATCACTG---ATATCTCTTCTCCACTGTATCAATCGGCTAT 24254
Qy 54 SerCysillelleTyrlleAla---GlyAlaSerGlnPheValilleThrAlaMetLeuAla 72
Db 24255 TG-AAGCTTGTGATGATGAGGAGCTTCTCATGCTGCTGCTTTC----- 24298
Qy 73 GlySerSerLeuTrp-----lleAlaAlaLeuThrValMetAlaMetAspVal 88
Db 24299 ---AGTCCATCTGCTCATTTAAGTCTTCTACACTGCTTATTCTAGTTATTCATTG 24355
Qy 89 ArgHisValLeuTyrlleGlyProSerLeuArgSerArgillelleGlnArgLeuGlnLysSer 108
Db 24356 TATCATCTTTTTCAGAGTTTATTAGCTTCTTGTGATGGGTAGAGATGCTCTTTAGCT 24415
Qy 109 LysThrAlaLeuTrpAlaPheGlyLeuThr-AspGluValPheAlaAlaAlaThrAlaL 128
Db 24416 CGGAGAAGTTGTTATTA-----CTGACCTCTGAAGCCTACTTCTGTCACATTGTC 24469
Qy 128 sLeuValArgAsnAsnArgArgTrpSerGluAsnTrpMetilleGlylleAlaPheSer 148
Db 24470 ACTCAAT-----CTTCATCCAGTTTGTCTTTCTTTG 24499
Qy 148 rTrpSerSer-----TrpValPheGlyThr---ValleGlyAlaPheSerGl 163
Db 24500 CTGGCAAGGAGCTGTGATCTTTTGGATAAGAGCACTCTGCTTTTGGAAATTTTACGCT 24559
Qy 163 ySerGlyLeuLeuGln-----GlyTyrlleProAlaValGluAlaAlaLeuGlyPh 179
Db 24560 CTCTGCTGCTTCTTCCCATCTTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 24619
Qy 179 eMetLeuProAlaLeuPheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLe 199
Db 24620 TGGTGTAGATGTCTTCTTGTGATGTATTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 24679
Qy 199 uCysValThrAlaAlaLeuValGlyAlaLeuAlaGlyValThrLeuPheSerlleProVa 219
Db 24680 TCTAACAGTCGCGCCTCTCAGCTGCGAGGTCTGTGGAGTTTGTGGAGTTCCTCCAGA 24739
Qy 219 lAlaileLeuAlaGlylleValCys---GlyCysLeuThrAlaLeuilleGlnAla----- 236
Db 24740 CCCTGTTTGCCTGGGTATACACAGTGGAGCTGCAGACCGCAATATTGCTGCTGATT 24799
Qy 237 ----PheTrpGlnGlyAlaPro 242
```

Db 24800 TTTCTTCTGGAAGCTTCATCCC 24821
RESULT 15
US-11-121-086-3
; Sequence 3, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121.086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 168516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-3

Alignment Scores:
Pred. No.: 8.08e+03 Length: 168516
Score: 89.50 Matches: 58
Percent Similarity: 35.6% Conservative: 32
Best Local Similarity: 22.9% Mismatches: 104
Query Match: 7.2% Indels: 60
DB: 12 Gaps: 8

US-10-073-293A-4 (1-245) x US-11-121-086-3 (1-168516)

Qy 6 ProGlnProAlaProGlySerAlaThrPheMetGluGlyCysLysAspSerLeuProIle 25
Db 9486 CCGCACCTTCCCTAGCGCGCCACCTGCTCCCTCCCGTACGGCTGTAGTCC 9545
Qy 26 ValIleSerTyrlleProVal----- 32
Db 9546 TCTGTGACCCCTGTTCACATCAGCTGACATGGAAGTTGGCCTGTCCCAACAGCAGGC 9605
Qy 33 ---AlaPheAlaPheGlyLeuAsnAlaThrArgLeuGlyPheSerProLeuGluSerVal 51
Db 9606 GCTGCTGGGGCCCAAGGCGCTTCTGTCAGGAGGC--GGGTTCCCGACCTCTGAT----- 9657
Qy 52 PhePheSerCysIleIleTyrlleAlaGlyAlaSerGlnPheValIleThrAlaMetLeuAla 71
Db 9658 -----CCCCCAGCAACGTCCTCCATGTAACTGGTTCTCTGTA 9696
Qy 72 AlaGlySerSerLeu-----TrrIleAlaAlaLeuThrVal 83
Db 9697 CAGGCACTGGCTTGTGGGCGCGCCAGGCTCTCCCATGGCTGGAGGTGCTCACTGTC 9756
Qy 84 MetAlaMetAspValArgHisValLeuTyrlleGlyProSerLeuArgSerArgIleIleGln 103
Db 9757 TGTGAGGCGCTCGCCAGCGCCCTGTGTCCAGAGCCTTCCCTGGGCTTCGGGGTCCCTCCA 9816
Qy 104 ArgLeuGlnLysSerLys--ThrAlaLeuTrpAlaPheGlyLeuThrAspGluValPheAl 123
Db 9817 GGTCTGGCCCGCAGGCGCCCTGTCTGTGGTGTGGCCATCTGCACCCCATCCCCA-----GC 9870
Qy 123 aAlaAlaThrAlaLysLeuValArgAsnAsnArgTrpSerGluAsnTrpMetIleG1 143
Db 9871 CACATCCTTGGCGCGCAGCTGTGGAGGATCTGAGGCGCTGGGGTAATGACTGCATGGCCAGG 9930
Qy 143 yIleAlaPheSerSerTrpSerSerTrp-----ValPheGlyTh 156
Db 9931 GTCAGGGTTGTCATCTGACGGCTCGACGAGGTGAACGCCCGCTCTCAGTCTTCCCTTC 9990
Qy 156 rValIleGlyAlaPheSerGlySerGlyLeuLeuGlnGlyTyrlleProAlaValGluAlaAl 176
Db 9991 TGTCTGTGGCGCTCACCGGTGGGCCCTCCCTGAGTGGGTGCTGCCACGCTG---GCCGG 10047

Search completed: February 16, 2006, 18:26:38
Job time : 705.857 secs

Qy 176 aLeuGlyPhe-----MetLeuPr 182
Db 10048 GGTGGATGGCAAAAGCGTGTGTGTGTGACAGGACGGCCCGAGGTTCACGTGGTCCCC 10107
Qy 182 oAlaLeu-----PheMetSerPheLeuLeuAlaSerPheGlnArgLysGlnSerLe 199
Db 10108 TGTGTTACAGGGCTTCGGAGGCCAGCTACCTGGACCGCAGCAGCAGCTGCAGGCCGTCT 10167
Qy 199 uCyeValThrAlaAlaLeuValGlyAlaLeuAlaGly 211
Db 10168 GTGCAGCACCATGTGGAAGGTGGCGGGGCTTCGGGA 10204

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 12:15:47 ; Search time 47.0815 Seconds
(without alignments)
1035.886 Million cell updates/sec

Title: US-10-073-293A-6
Perfect score: 552
Sequence: 1 MSTEVLLGLLVGVNYCFR.....IPTLLSALAYGLAWKNVAIL 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	552	100.0	111	5	ABP57757 E. coli L
2	552	100.0	111	9	ADZ39953 E.coli b2
3	552	100.0	111	9	ADZ36197 Escherich
4	403	73.0	123	7	ABO63701 Klebsiell
5	231	41.8	74	4	RAM84790 Human imm
6	182.5	33.1	122	6	ADA35577 Acinetoba
7	78	14.1	530	6	ABU35859 Protein e
8	77.5	14.0	442	6	ABU35020 Protein e
9	77	13.9	308	8	ADN20876 Bacterial
10	76.5	13.9	494	2	AAY27289 Glucose t
11	76.5	13.9	500	8	ADS23412 Bacterial
12	76	13.8	463	3	AAG04739 Arabidops
13	76	13.8	502	3	AAG04738 Arabidops
14	76	13.8	509	5	ABU65197 Human NOV
15	76	13.8	509	8	ADN62045 Human nov
16	76	13.8	561	3	AAG04737 Arabidops
17	76	13.8	650	6	ABU42744 Protein e
18	76	13.8	655	5	ABP40686 Staphyloc
19	76	13.8	655	8	ADN07960 Staphyloc
20	75.5	13.7	201	6	ADB11930 Alloiococ
21	75.5	13.7	225	6	ADB11928 Alloiococ
22	75.5	13.7	253	6	ADB11926 Alloiococ
23	75.5	13.7	254	6	ADB11924 Alloiococ
24	75.5	13.7	283	6	ADB11922 Alloiococ

25	75.5	13.7	326	6	ADB11920
26	75.5	13.7	327	6	ADB11918
27	75.5	13.7	497	8	ADQ66449 Novel hum
28	75.5	13.7	497	9	ADX87332 Human glu
29	75.5	13.7	520	8	ADQ66738 Novel hum
30	74	13.4	202	4	AA878985
31	74	13.4	202	4	AA878984
32	74	13.4	472	5	ABU05729
33	74	13.4	791	4	AAG90776
34	74	13.4	791	7	ADL65603
35	73.5	13.3	156	2	AAR90770
36	73.5	13.3	156	2	AAW46893
37	73.5	13.3	156	6	ABR58649
38	73.5	13.3	156	8	ABM80157
39	73.5	13.3	156	9	ADV70187
40	73.5	13.3	332	7	ABO62813
41	73	13.2	269	5	ABO47774
42	73	13.2	298	8	ADV88288
43	73	13.2	298	8	ADV81710
44	73	13.2	298	8	ADV79541
45	73	13.2	319	5	ABP29846

ALIGNMENTS

RESULT 1
ABP57757
ID ABP57757 standard; protein; 111 AA.
XX
AC ABP57757;
XX
DT 29-JAN-2003 (first entry)
XX
DE E. coli L-amino acid producing protein #2.
XX
KW L-amino acid; E. coli.
XX
OS Escherichia coli.
XX
PN EP1239041-A2.
XX
PD 11-SEP-2002.
XX
PF 13-FEB-2002; 2002EP-00003335.
XX
PR 13-FEB-2001; 2001RU-00103865.
PR 26-FEB-2001; 2001RU-00104998.
PR 26-FEB-2001; 2001RU-00104999.
PR 28-JUN-2001; 2001RU-00117632.
PR 28-JUN-2001; 2001RU-00117633.
XX (AJIN) AJINOMOTO CO INC.

Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Guseyatiner MM;
WPI; 2002-699856/76.
N-PSDB; ABV75657.
Novel L-amino acid producing Escherichia bacterium, is modified to enhance L-amino acid production by enhancing the activities of protein capable of making bacterium to have enhanced resistance to L-amino acids.
Disclosure; Page 24-25; 33pp; English.
The invention relates to a novel L-amino acid producing bacterium of genus Escherichia, modified to enhance L-amino acid production by enhancing the activities of a protein. The novel bacterium is useful for producing L-amino acid e.g. L-Thr, L-Pro, L-Leu, L-Met and L-Arg, by cultivating the bacterium in a culture medium and collecting L-amino acid to be produced and accumulated from the culture medium. The present sequence represents a protein of the invention which causes increased L-amino acid production in E. coli

XX SQ Sequence 111 AA;

Query Match 100.0%; Score 552; DB 5; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSYEVLIGLLGVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASCALLVNSTA 60
 |||||
 Db 1 MSYEVLIGLLGVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASCALLVNSTA 60
 |||||

Qy 61 PEVMDTRFRFVPTLVGFAVLGASFYKTRSIIPILLSSALAYGLAWKVMII 111
 |||||
 Db 61 PEVMDTRFRFVPTLVGFAVLGASFYKTRSIIPILLSSALAYGLAWKVMII 111
 |||||

RESULT 2
 ADZ39953
 ID ADZ39953 standard; protein; 111 AA.

XX AC ADZ39953;
 XX DT 14-JUL-2005 (first entry)
 XX DE E.coli b2683 protein for improved amino acid production in bacteria.
 XX KW amino acid production; fermentation; transport protein.
 XX OS Escherichia coli.
 XX PN EP1526179-A1.
 XX PD 27-APR-2005.

XX PF 13-FEB-2002; 2004EP-00028876.
 XX PR 13-FEB-2001; 2001RU-00103865.
 XX PR 26-FEB-2001; 2001RU-00104998.
 XX PR 26-FEB-2001; 2001RU-00104999.
 XX PR 28-JUN-2001; 2001RU-00117632.
 XX PR 28-JUN-2001; 2001RU-00117633.
 XX PR 13-FEB-2002; 2002EP-00003335.
 XX PA (AJIN) AJINOMOTO CO INC.
 XX PI Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Gussyatiner MM;
 XX DR WPI; 2005-308098/32.
 XX DR N-PSDB; ADZ39952.

XX PT New L-amino acid producing bacterium belonging to the genus Escherichia,
 PT useful for producing L-amino acids, e.g. L-threonine, L-valine, L-
 PT proline, L-methionine, or L-arginine.
 XX PS Disclosure; SEQ ID NO 6; 35pp; English.

XX CC The invention relates to an L-amino acid producing bacterium belonging to
 CC the genus Escherichia, where the bacterium has been modified so that the
 CC L-amino acid production by the bacterium is enhanced by enhancing
 CC activities of proteins by transformation of the bacterium with DNA coding
 CC for protein or by alteration of an expression regulation sequence of the
 CC DNA on the chromosome of the bacterium. The L-amino acid producing
 CC bacterium is useful for producing L-amino acids by fermentation,
 CC including L-threonine, L-valine, L-proline, L-methionine, or L-arginine.
 CC The genes are useful for improving L-amino acid productivity. This
 CC sequence corresponds to the b2683 protein, one of the L-amino acid
 CC biosynthesis pathway proteins. The protein is a putative transport
 CC protein within the biosynthesis pathway. Enhancement of the expression of
 CC this protein increases production on L-threonine, L-valine, L-leucine, L-
 CC proline and L-methionine.

XX SQ Sequence 111 AA;

Query Match 100.0%; Score 552; DB 9; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSYEVLIGLLGVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASCALLVNSTA 60
 |||||
 Db 1 MSYEVLIGLLGVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASCALLVNSTA 60
 |||||

Qy 61 PEVMDTRFRFVPTLVGFAVLGASFYKTRSIIPILLSSALAYGLAWKVMII 111
 |||||
 Db 61 PEVMDTRFRFVPTLVGFAVLGASFYKTRSIIPILLSSALAYGLAWKVMII 111
 |||||

RESULT 3
 ADZ36197
 ID ADZ36197 standard; protein; 111 AA.

XX AC ADZ36197;
 XX DT 14-JUL-2005 (first entry)
 XX DE Escherichia coli b2683 protein SEQ ID NO:6.
 XX KW amino acid production.
 XX OS Escherichia coli.
 XX PN EP1526181-A1.
 XX PD 27-APR-2005.

XX PF 13-FEB-2002; 2004EP-00028877.
 XX PR 13-FEB-2001; 2001RU-00103865.
 XX PR 26-FEB-2001; 2001RU-00104998.
 XX PR 26-FEB-2001; 2001RU-00104999.
 XX PR 28-JUN-2001; 2001RU-00117632.
 XX PR 28-JUN-2001; 2001RU-00117633.
 XX PR 13-FEB-2002; 2002EP-00003335.
 XX PA (AJIN) AJINOMOTO CO INC.
 XX PI Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Gussyatiner MM;
 XX DR WPI; 2005-317157/33.
 XX DR N-PSDB; ADZ36196.

XX PT New modified L-amino acid producing bacterium useful to enhance the
 PT production of L-amino acid by enhancing the activities of proteins in a
 PT cell of bacterium.
 XX PS Disclosure; SEQ ID NO 6; 35pp; English.

XX CC The invention relates to an L-amino acid (A) producing bacterium (1)
 CC (belonging to the genus Escherichia), which is modified to enhance the
 CC production of (A) by enhancing the activities of proteins (G) or (H) in a
 CC cell of (1). Also described is a method for producing (A) comprising
 CC cultivating the bacterium in a culture medium and collecting the produced
 CC and accumulated L-amino acid. The modified bacterium has the ability to
 CC grow on a minimal medium containing L-amino acid or its analog in a
 CC minimal concentration and ability to grow faster on a medium containing L
 CC -amino acid or its analog than the unmodified strain or the wild type
 CC strain, or the parental strain of the bacterium. The present sequence
 CC represents the E. coli b2683 protein which has L-amino acid excretion
 CC activity.

XX SQ Sequence 111 AA;

Query Match 100.0%; Score 552; DB 9; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSYEVLIGLLGVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASCALLVNSTA 60

Db 1 MSYEVLGLLVGVVAVYCFRLPLRLRVGNARPTKRGAVGILLDTIGTASTCALLVVSTA 60
Qy 61 PEVMDTRRRFVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKVMII 111
Db 61 PEVMDTRRRFVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKVMII 111

RESULT 4
ABO63701
ID ABO63701 standard; protein; 123 AA.
XX ABO63701;
AC ABO63701;
XX 29-JUL-2004 (first entry)
XX Klebsiella pneumoniae polypeptide seqid 10218.
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX Klebsiella pneumoniae.
OS US6610836-B1.
PN 26-AUG-2003.
XX 27-JAN-2000; 2000US-00489039.
XX 29-JAN-1999; 99US-0117747P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Breton GL, Osborne M;
XX WPT; 2003-895346/82.
DR N-PSDB; ACH97252.
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX Disclosure; SEQ ID NO 10218; 932pp; English.
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX Sequence 123 AA;

Query Match 73.0%; Score 403; DB 7; Length 123;
Best Local Similarity 72.0%; Pred. No. 6.6e-42;
Matches 77; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

Qy 1 MSYEVLGLLVGVVAVYCFRLPLRLRVGNARPTKRGAVGILLDTIGTASTCALLVVSTA 60
Db 13 MNSYVLLGVVGVVAVYCFRLPLRLRVGNARPTKRGAVGILLDTIGTASTCALLVVSTA 72
Qy 61 PEVMDTRRRFVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKV 107
Db 73 PEILADARLAPLTGFLVGLGAPFWKTRSIIPITLLSAPAYGLAWKI 119

RESULT 5
AAM84790
ID AAM84790 standard; protein; 74 AA.
XX AAM84790;
XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen SEQ ID NO:12383.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX Homo sapiens.
OS WO200157182-A2.
PN 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US001354.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226988P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231422P.
PR 08-SEP-2000; 2000US-0231423P.
PR 08-SEP-2000; 2000US-0231443P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.

Qy 64 MHDTRRFVPTLVGFAVLGASFYKTRSTIIPTLLSALAY 101
 Db 149 AH-----APALIPRTGL-ANFYRGFS---PTLLGMLPY 177

RESULT 10
 AAY27289
 ID AAY27289 standard; protein; 494 AA.

AC AAY27289;
 DT 05-NOV-1999 (first entry)
 XX Glucose transporter protein GLUT3.
 XX Human; glucose transporter; GLUTX; hexose transport; gene therapy;
 KW chromosome mapping; GLUT3.
 KW Homo sapiens.
 XX US5942398-A.
 XX 24-AUG-1999.
 XX 26-FEB-1998; 98US-00031392.
 XX 26-FEB-1998; 98US-00031392.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Tartaglia LA, Weng X;
 XX WPI; 1999-526192/44.

XX A novel mammalian transporter protein and the gene encoding it, are
 PT useful in the diagnosis and treatment of disorders associated with
 PT aberrant sugar transport.

XX Example; Fig 3; 48pp; English.

XX The invention relates to a human glucose transporter (GLUTX) protein. The
 CC protein can be expressed by standard recombinant methodology. GLUTX
 CC nucleic acids are useful as hybridization probes for detecting the
 CC presence of GLUTX DNA in a sample, useful for diagnosing conditions
 CC associated with aberrant expression levels of GLUTX. The GLUTX gene is
 CC also useful as a therapeutic agent for regulating translation of GLUTX
 CC mRNA, and for treatment of disorders associated with aberrant expression
 CC of GLUTX and aberrant hexose transport. It is useful for generating GLUTX
 CC specific antibodies, identifying agonists and antagonists of GLUTX, and
 CC identifying nucleic acids in other species encoding nucleic acids
 CC homologous to GLUTX. GLUTX nucleic acids are also useful for identifying
 CC the chromosomal location of GLUTX, and as tissue specific markers.
 CC Sequences AAY27287-291 represent different glucose transporter proteins
 CC which were compared with human GLUTX

XX Sequence 494 AA;
 Query Match 13.9%; Score 76.5; DB 2; Length 494;
 Best Local Similarity 34.1%; Pred. No. 2;
 Matches 29; Conservative 17; Mismatches 30; Indels 9; Gaps 7;

Qy 2 SYEVLLG-LLVGV-ANVCYRLRLRVGNARPTK-RGAVGILLDTIGIASICALVWS 58
 Db 116 SVEMLILGRLLIGLFCGLCTGFVP--WYGEISFTALRGAFG-TLNQGLI--VIGILVAQ 170
 Qy 59 T-APEVMDTRRRFVPTLVGFAVLGA 82
 Db 171 IFGLKVILGTEDLWPLLGLFTILPA 195

RESULT 11
 ADS23412
 ID ADS23412 standard; protein; 500 AA.

XX ADS23412;
 XX 02-DEC-2004 (first entry)
 XX Bacterial polypeptide #12445.
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX Bacteria.
 XX US2003233675-A1.
 XX 18-DEC-2003.
 XX 20-FEB-2003; 2003US-00369493.
 XX 21-FEB-2002; 2002US-0360039P.
 XX (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.
 XX (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 12445; 122pp; English.
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 500 AA;
 Query Match 13.9%; Score 76.5; DB 8; Length 500;
 Best Local Similarity 25.2%; Pred. No. 2.1;
 Matches 33; Conservative 20; Mismatches 45; Indels 33; Gaps 5;
 Qy 9 GLLVGAVNYCFYRLRLRVGNARPTKRGAVGILLDTIGIASICALVWSTAPEVMDTR 68

```
Db      20 GLFVGIGAYLRDAGFLSLLLGY---LWGIAFILPINLUSVCEMCAYLPIRGS--IFELAA 74
Qy      69 RFVPTLVGFVAVLGASFYK---TRS-----IIPTLLSALAYGL----- 103
Db      75 RYVDPAFGFMKXGFSAPRVQGRSDNIRLGDWVYFAGLMLVCTEYSAVAFIMDYWQIDV 134
Qy      104 ---AWKUNAI 111
Db      135 NPAAWVANAMV 145

RESULT 12
AAG04739
ID  AAG04739 standard; protein; 463 AA.
XX
AC  AAG04739;
XX
DT  17-OCT-2000 (first entry)
XX
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 879.
XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX
OS  Arabidopsis thaliana.
XX
PN  EP1033405-A2.
XX
PD  06-SEP-2000.
XX
PF  25-FEB-2000; 2000EP-00301439.
XX
PR  25-FEB-1999; 99US-0121825P.
PR  05-MAR-1999; 99US-01231180P.
PR  09-MAR-1999; 99US-0123548P.
PR  23-MAR-1999; 99US-0125788P.
PR  25-MAR-1999; 99US-0126264P.
PR  29-MAR-1999; 99US-0126785P.
PR  01-APR-1999; 99US-0127462P.
PR  06-APR-1999; 99US-0128234P.
PR  08-APR-1999; 99US-0128714P.
PR  16-APR-1999; 99US-0129845P.
PR  19-APR-1999; 99US-0130077P.
PR  21-APR-1999; 99US-0130449P.
PR  23-APR-1999; 99US-0130510P.
PR  28-APR-1999; 99US-0130891P.
PR  30-APR-1999; 99US-0131449P.
PR  30-APR-1999; 99US-0132048P.
PR  30-APR-1999; 99US-0132407P.
PR  04-MAY-1999; 99US-0132484P.
PR  05-MAY-1999; 99US-0132485P.
PR  06-MAY-1999; 99US-0132486P.
PR  06-MAY-1999; 99US-0132487P.
PR  07-MAY-1999; 99US-0132863P.
PR  11-MAY-1999; 99US-0134258P.
PR  14-MAY-1999; 99US-0134218P.
PR  14-MAY-1999; 99US-0134219P.
PR  14-MAY-1999; 99US-0134221P.
PR  14-MAY-1999; 99US-0134370P.
PR  18-MAY-1999; 99US-0134768P.
PR  19-MAY-1999; 99US-0134941P.
PR  20-MAY-1999; 99US-0135124P.
PR  21-MAY-1999; 99US-0135353P.
PR  24-MAY-1999; 99US-0135629P.
PR  25-MAY-1999; 99US-0136021P.
PR  27-MAY-1999; 99US-0136382P.
PR  28-MAY-1999; 99US-0136782P.
PR  01-JUN-1999; 99US-0137222P.
PR  03-JUN-1999; 99US-0137528P.
PR  04-JUN-1999; 99US-0137502P.
PR  07-JUN-1999; 99US-0137724P.
PR  08-JUN-1999; 99US-0138094P.
PR  10-JUN-1999; 99US-0138540P.
PR  10-JUN-1999; 99US-0138847P.
PR  14-JUN-1999; 99US-0139119P.
PR  16-JUN-1999; 99US-0139452P.
PR  16-JUN-1999; 99US-0139453P.
PR  17-JUN-1999; 99US-0139492P.
PR  18-JUN-1999; 99US-0139454P.
PR  18-JUN-1999; 99US-0139455P.
PR  18-JUN-1999; 99US-0139456P.
PR  18-JUN-1999; 99US-0139457P.
PR  18-JUN-1999; 99US-0139458P.
PR  18-JUN-1999; 99US-0139459P.
PR  18-JUN-1999; 99US-0139460P.
PR  18-JUN-1999; 99US-0139461P.
PR  18-JUN-1999; 99US-0139462P.
PR  18-JUN-1999; 99US-0139463P.
PR  18-JUN-1999; 99US-0139750P.
PR  18-JUN-1999; 99US-0139763P.
PR  21-JUN-1999; 99US-0139817P.
PR  22-JUN-1999; 99US-0139899P.
PR  23-JUN-1999; 99US-0140353P.
PR  23-JUN-1999; 99US-0140354P.
PR  24-JUN-1999; 99US-0140695P.
PR  28-JUN-1999; 99US-0140823P.
PR  29-JUN-1999; 99US-0140991P.
PR  30-JUN-1999; 99US-0141287P.
PR  01-JUL-1999; 99US-0141842P.
PR  01-JUL-1999; 99US-0142154P.
PR  02-JUL-1999; 99US-0142055P.
PR  06-JUL-1999; 99US-0142390P.
PR  08-JUL-1999; 99US-0142803P.
PR  09-JUL-1999; 99US-0142920P.
PR  12-JUL-1999; 99US-0142977P.
PR  13-JUL-1999; 99US-0143542P.
PR  14-JUL-1999; 99US-0143624P.
PR  15-JUL-1999; 99US-0144005P.
PR  16-JUL-1999; 99US-0144085P.
PR  16-JUL-1999; 99US-0144086P.
PR  19-JUL-1999; 99US-0144325P.
PR  19-JUL-1999; 99US-0144331P.
PR  19-JUL-1999; 99US-0144332P.
PR  19-JUL-1999; 99US-0144333P.
PR  19-JUL-1999; 99US-0144334P.
PR  19-JUL-1999; 99US-0144335P.
PR  20-JUL-1999; 99US-0144352P.
PR  20-JUL-1999; 99US-0144632P.
PR  20-JUL-1999; 99US-0144884P.
PR  21-JUL-1999; 99US-0144814P.
PR  21-JUL-1999; 99US-0145086P.
PR  21-JUL-1999; 99US-0145088P.
PR  22-JUL-1999; 99US-0145085P.
PR  22-JUL-1999; 99US-0145087P.
PR  22-JUL-1999; 99US-0145089P.
PR  22-JUL-1999; 99US-0145192P.
PR  23-JUL-1999; 99US-0145145P.
PR  23-JUL-1999; 99US-0145216P.
PR  23-JUL-1999; 99US-0145224P.
PR  26-JUL-1999; 99US-0145276P.
PR  27-JUL-1999; 99US-0145913P.
PR  27-JUL-1999; 99US-0145918P.
PR  27-JUL-1999; 99US-0145919P.
PR  28-JUL-1999; 99US-0145951P.
PR  02-AUG-1999; 99US-0146386P.
PR  02-AUG-1999; 99US-0146388P.
PR  02-AUG-1999; 99US-0146389P.
PR  03-AUG-1999; 99US-0147038P.
PR  04-AUG-1999; 99US-0147204P.
PR  04-AUG-1999; 99US-0147302P.
PR  05-AUG-1999; 99US-0147192P.
PR  05-AUG-1999; 99US-0147260P.
PR  06-AUG-1999; 99US-0147303P.
PR  06-AUG-1999; 99US-0147416P.
```

PR	09-AUG-1999;	99US-0147493P.	QY	6	LLGLLVGVANYCFRYLPLRLRVGNARPTKRGAVGILLD-----TIGIASICAL- 54
PR	09-AUG-1999;	99US-0147935P.	Db	4	WIGLVGVPTV---YL-----AGSLVDLGMWQGIATVVTANLLLV 44
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.	QY	55	-LVVSTAPEVNHDTTRFVPTLVGFVNLGASFYKTRSIITLLSALAYGLAW 105
PR	13-AUG-1999;	99US-0148684P.	Db	45	PLVLTAPQPGTLYG-----ISFPVLARSFGIRGAHIPTLLRALV-GCGW 87
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151085P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157533P.			
PR	07-OCT-1999;	99US-0157865P.			
PR	08-OCT-1999;	99US-0158029P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159594P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.			
PR	22-OCT-1999;	99US-0160980P.			
PR	22-OCT-1999;	99US-0160981P.			
PR	22-OCT-1999;	99US-0160989P.			
PR	25-OCT-1999;	99US-0161404P.			
PR	25-OCT-1999;	99US-0161405P.			
PR	25-OCT-1999;	99US-0161406P.			
PR	26-OCT-1999;	99US-0161359P.			
PR	26-OCT-1999;	99US-0161360P.			
PR	26-OCT-1999;	99US-0161361P.			
PR	28-OCT-1999;	99US-0161920P.			
PR	28-OCT-1999;	99US-0161922P.			
PR	28-OCT-1999;	99US-0161931P.			
PR	29-OCT-1999;	99US-0162142P.			

Query Match 13.8%; Score 76; DB 3; Length 463;
 Best Local Similarity 28.6%; Pred. No. 2.2;
 Matches 32; Conservative 12; Mismatches 28; Indels 40; Gaps 6;

RESULT 13

AAG04738

ID AAG04738 standard; protein; 502 AA.

XX

AC AAG04738;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 878.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

25-FEB-2000; 2000EP-00301439.

XX

25-FEB-1999; 99US-0121825P.

PR

05-MAR-1999; 99US-0123180P.

PR

03-MAR-1999; 99US-0123548P.

PR

23-MAR-1999; 99US-0125788P.

PR

25-MAR-1999; 99US-0126264P.

PR

01-APR-1999; 99US-0127462P.

PR

06-APR-1999; 99US-0128234P.

PR

08-APR-1999; 99US-0128714P.

PR

16-APR-1999; 99US-0129845P.

PR

19-APR-1999; 99US-0130077P.

PR

21-APR-1999; 99US-0130449P.

PR

23-APR-1999; 99US-0130891P.

PR

28-APR-1999; 99US-0131449P.

PR

30-APR-1999; 99US-0132048P.

PR

30-APR-1999; 99US-0132407P.

PR

04-MAY-1999; 99US-0132484P.

PR

05-MAY-1999; 99US-0132485P.

PR

06-MAY-1999; 99US-0132486P.

PR

07-MAY-1999; 99US-0132863P.

PR

11-MAY-1999; 99US-0134256P.

PR

14-MAY-1999; 99US-0134218P.

PR

14-MAY-1999; 99US-0134219P.

PR

14-MAY-1999; 99US-0134221P.

PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142300P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 22-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 13.8%; Score 76; DB 3; Length 502;

Best Local Similarity 28.6%; Pred. No. 2.4; Mismatches 28; Indels 40; Gaps 6;

Matches 32; Conservative 12;

Oy 6 LLLGLLVGVANYCFRFLPLRLRVGNARPTKRGAVGILLD-----TTGIAISICAL- 54

Db 43 LWWGLLVGVPTY---YL-----AGSLVDLGMWVQGIATVVTANLILV 83

Qy 55 -LVVSTAPEVMDTRRFPVTLVGVAVIGASPKYKTRSIIPITLSALAYGLAW 105
 Db 84 PLVLTQAQGTLYG-----ISFPVLARSFGIRGAHIPTLLRALV-GCGW 126

RESULT 14
 ABU65197
 ID ABU65197 standard; protein; 509 AA.

AC ABU65197;
 XX 20-MAY-2003 (first entry)
 XX Human NOV112a protein.
 XX NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KW human.

XX Homo sapiens.

XX WO200272757-A2.

XX 19-SEP-2002.

XX 08-MAR-2002; 2002WO-US006908.

XX 08-MAR-2001; 2001US-0274101P.
 PR 08-MAR-2001; 2001US-0274194P.
 PR 08-MAR-2001; 2001US-0274281P.
 PR 08-MAR-2001; 2001US-0274322P.
 PR 09-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275235P.
 PR 13-MAR-2001; 2001US-0275578P.
 PR 13-MAR-2001; 2001US-0275579P.
 PR 13-MAR-2001; 2001US-0275601P.
 PR 14-MAR-2001; 2001US-0276000P.
 PR 16-MAR-2001; 2001US-0276776P.
 PR 19-MAR-2001; 2001US-0276994P.
 PR 20-MAR-2001; 2001US-0277239P.
 PR 20-MAR-2001; 2001US-0277321P.
 PR 20-MAR-2001; 2001US-0277327P.
 PR 21-MAR-2001; 2001US-0277791P.
 PR 22-MAR-2001; 2001US-0277833P.
 PR 23-MAR-2001; 2001US-0278152P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0278999P.
 PR 27-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0279388P.
 PR 30-MAR-2001; 2001US-0279995P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 02-APR-2001; 2001US-0280900P.
 PR 04-APR-2001; 2001US-0281194P.
 PR 13-APR-2001; 2001US-0283675P.
 PR 30-APR-2001; 2001US-0287424P.
 PR 02-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 03-MAY-2001; 2001US-0288528P.
 PR 15-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 16-MAY-2001; 2001US-0291240P.
 PR 30-MAY-2001; 2001US-0294485P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.

PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 14-NOV-2001; 2001US-0333272P.
 PR 21-NOV-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 04-DEC-2001; 2001US-0338092P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 07-MAR-2002; 2002US-00092900.
 XX
 PA (CURA-) CURAGEN CORP.

XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
 PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernov V;
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
 PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
 PI Lepley DM, Rieser DK;
 XX

XX WPI; 2002-723332/78.
 DR N-PSDB; ABX97164.

XX NOVX polypeptides and polynucleotides, useful for preventing or treating
 PT a disorder associated with aberrant NOVX expression or activity e.g.,
 PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
 PT asthma.

XX Claim 1; Page 428; 1103pp; English.

XX This invention describes novel human NOVX polypeptides which have
 CC cytotatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
 CC activity. Pharmaceutical compositions comprising the NOVX proteins or
 CC nucleic acid molecules or NOVX antibodies are useful for preventing or
 CC treating a disorder associated with aberrant NOVX expression or activity
 CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
 CC asthma. The products of the invention can be used for gene therapy or in
 CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
 CC ABX97008-ABX97185

XX Sequence 509 AA;

Query Match 13.8%; Score 76; DB 5; Length 509;
 Best Local Similarity 32.1%; Pred. No. 2.4;
 Matches 27; Conservative 17; Mismatches 34; Indels 6; Gaps 5;

Qy 2 SYEVLLLG-LLVGV-ANYCFRYLRLRVGNARPTK-RGAVGILLDTIGTASICALVWS 58
 Db 119 SVEMLILGRVLGLFCGLCTGFPV--MYIGEISPTALRGAFG-TLNQLGIVIGILVAQVI 175
 Qy 59 TAEVMDTRRFPVTLVGVAVLCA 82
 Db 176 FGLELIGSEELMPVLLGFTILPA 199

RESULT 15
 ADN62045
 ID ADN62045 standard; protein; 509 AA.

XX ADN62045;

XX 01-JUL-2004 (first entry)

XX Human novel protein NOV112a.

KW Human; NOVX; diabetes; obesity; infectious disease; anorexia;
KW cancer-associated cachexia; cancer; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; chronic disease.

XX Homo sapiens.

XX US2004043382-A1.

PN 04-MAR-2004.

XX 07-MAR-2002; 2002US-00092900.

XX 08-MAR-2001; 2001US-0274191P.

PR 08-MAR-2001; 2001US-0274194P.

PR 08-MAR-2001; 2001US-0274281P.

PR 08-MAR-2001; 2001US-0274322P.

PR 09-MAR-2001; 2001US-0274849P.

PR 12-MAR-2001; 2001US-0275235P.

PR 13-MAR-2001; 2001US-0275578P.

PR 13-MAR-2001; 2001US-0275579P.

PR 13-MAR-2001; 2001US-0275601P.

PR 14-MAR-2001; 2001US-0276000P.

PR 16-MAR-2001; 2001US-0276776P.

PR 19-MAR-2001; 2001US-0276994P.

PR 20-MAR-2001; 2001US-0277239P.

PR 20-MAR-2001; 2001US-0277321P.

PR 20-MAR-2001; 2001US-0277327P.

PR 20-MAR-2001; 2001US-0277338P.

PR 21-MAR-2001; 2001US-0277791P.

PR 22-MAR-2001; 2001US-0277833P.

PR 23-MAR-2001; 2001US-0278152P.

PR 26-MAR-2001; 2001US-0278894P.

PR 27-MAR-2001; 2001US-0278999P.

PR 27-MAR-2001; 2001US-0279036P.

PR 28-MAR-2001; 2001US-0279344P.

PR 30-MAR-2001; 2001US-0279995P.

PR 30-MAR-2001; 2001US-0280233P.

PR 02-APR-2001; 2001US-0280802P.

PR 02-APR-2001; 2001US-0280802P.

PR 02-APR-2001; 2001US-0280802P.

PR 04-APR-2001; 2001US-0280900P.

PR 13-APR-2001; 2001US-0281444P.

PR 13-APR-2001; 2001US-0281675P.

PR 30-APR-2001; 2001US-0287424P.

PR 02-MAY-2001; 2001US-0288066P.

PR 03-MAY-2001; 2001US-0288342P.

PR 04-DEC-2001; 2001US-0337185P.

XX 03-JAN-2002; 2002US-0345705P.

PA (PADI/) PADIGARU M.

PA (SPYT/) SPYTEK K A.

PA (SHEN/) SHENOY S G.

PA (TAUP/) TAUPIER R J.

PA (PENA/) PENA C E A.

PA (LILL/) LI L.

PA (ZERH/) ZERHUSEN B D.

PA (GUSE/) GUSEV V Y.

PA (JIW/) JI W.

PA (GORM/) GORMAN L.

PA (MILL/) MILLER C E.

PA (KEKU/) KEKUDA R.

PA (PATT/) PATTURAJAN M.

PA (GANG/) GANGOLLI E A.

PA (VERN/) VERNET C A M.

PA (GUOX/) GUO X S.

PA (TCHE/) TCHERNEV V T.

PA (FERN/) FERNANDES E R.

PA (CASM/) CASMAN S J.

PA (MALY/) MALYANKAR U M.

PA (GERL/) GERLACH V.

PA (LIUY/) LIU Y.

PA (ANDE/) ANDERSON D W.

PA (SPAD/) SPADERNA S K.

PA (CATT/) CATTERTON E.

PA (LEIT/) LEITE M W.

PA (ZHON/) ZHONG H.

PA (ALSO/) ALSOBROOK J P.

PA (LEPL/) LEPLEY D M.

PA (RIEG/) RIEGER D K.

PA (BURG/) BURGESS C E.

XX WPI; 2004-225693/21.

PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;

PI Zehusen BD, Gusev VV, Ji W, Gorman L, Miller CE, Kekuda R;

PI Patturajan M, Gangolli EA, Vernet CAM, Guo XS, Tchernev VT;

PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;

PI Anderson DM, Spaderna SK, Catterton E, Leite MW, Zhong H;

PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;

XX WPI; 2004-225693/21.

DR N-PSDB; ADM62044.

XX New NOVX polypeptides and nucleic acid molecules useful for diagnosing,

PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,

PT infection or obesity, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX Claim 1; SEQ ID NO 314; 786pp; English.

PS The invention relates to an isolated polypeptide (designated NOVX, or

XX NOV1-NOV127) comprising a sequence selected from 178 fully defined amino

CC acid sequences (and their mature forms, variants and fragments). Also

CC included are an isolated nucleic acid molecule encoding NOVX, a vector

CC comprising the nucleic acid, a cell comprising the vector, methods for

CC determining the presence or amount of the polypeptide or the nucleic acid

CC molecule in a sample, methods for determining the presence of or

CC predisposition to a disease associated with altered levels of expression

CC of the above polypeptide or nucleic acid molecule in a first mammalian

CC subject, a method for identifying an agent that binds to the above

CC polypeptide, a method for identifying a potential therapeutic agent for

CC use in the treatment of a pathology that is related to aberrant

CC expression or physiological interactions of the polypeptide, a method of

CC screening for a modulator of activity or of latency or predisposition to

CC a pathology associated with the polypeptide and a method for modulating

CC the activity of the polypeptide cited above. The composition and methods

CC are useful for diagnosing, preventing or treating diseases such as

CC diabetes, obesity, infectious diseases, anorexia, cancer-associated

CC cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or

CC Parkinson's disease, immune disorders, haematopoietic disorders,

CC dyslipidaemias, and other chronic diseases. These may also be used in

CC chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The polypeptides are also useful as vaccines. The
CC present sequence represents a NOVX protein of the invention.
XX
SQ Sequence 509 AA;
Query Match 13.8%; Score 76; DB 8; Length 509;
Best Local Similarity 32.1%; Pred. No. 2.4;
Matches 27; Conservative 17; Mismatches 34; Indels 6; Gaps 5;
Qy 2 SYEVLLG-LLVGV-ANVCFRYLPRLRLRVGNARPTK-RGAVGILLDTIGIASICALLVWS 58
Db 119 SVEMLILGRVLVIGLFCGLCTGFP--WYGEISPTALRGAPG-TINQIGIVIGILVAQVI 175
Qy 59 TAPEVMHDTRRFVPTLVGFAVLGA 82
Db 176 FGLEILGSBELWPVLLGFTILPA 199

Search completed: February 15, 2006, 12:18:34
Job time : 50.0815 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 12:18:47 ; Search time 14.9663 Seconds
(without alignments)
713.608 Million cell updates/sec

Title: US-10-073-293A-6
Perfect score: 552
Sequence: 1 MSYEVLGLLVGVNYCFR.....IPTLLSALAYGLAWKVMII 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	552	100.0	111	2 D65048	hypothetical prote
2	548	99.3	111	2 A91072	hypothetical prote
3	544	98.6	111	2 D85916	hypothetical prote
4	308.5	55.9	113	2 A90396	probable membrane
5	78.5	14.2	866	2 AG5272	hypothetical prote
6	78	14.1	530	2 G87022	hypothetical prote
7	76.5	13.9	309	2 B69556	phosphate ABC tran
8	76	13.8	599	2 T48383	uracil transporter
9	74.5	13.5	277	1 H71220	hypothetical prote
10	74.5	13.5	486	2 H84805	hypothetical prote
11	74	13.4	472	2 H70836	hypothetical prote
12	73.5	13.3	373	2 A49806	prv43 protein - su
13	73	13.2	269	2 AG1126	conserved membrane
14	73	13.2	269	2 AD1487	conserved membrane
15	73	13.2	459	2 S28025	light harvesting c
16	73	13.2	672	2 D85891	hydrogenase 4 memb
17	72.5	13.1	496	2 A31986	glucose transporte
18	72.5	13.1	637	2 T08530	tag protein - Ent
19	72.5	13.1	637	2 S22992	trac protein - Esc
20	72	13.0	481	2 B86285	hypothetical prote
21	72	13.0	660	1 S54746	cytochrome c-type
22	71.5	13.0	431	2 B96006	probable metabolit
23	71.5	13.0	1037	2 AE0816	probable efflux pu
24	71	12.9	388	2 AB0507	Na+/H+ antiporter
25	71	12.9	464	2 C75474	conserved hypotet
26	70.5	12.8	440	2 T50912	hypothetical prote
27	70.5	12.8	542	2 AF2587	MFS permease limpo
28	70.5	12.8	542	2 F97369	hypothetical prote
29	70	12.7	100	2 A69469	conserved hypotet

ALIGNMENTS

RESULT 1

D65048

hypothetical protein (emrR 5' region) - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C;Accession: D65048

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: D65048

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-111 <BLAT>

A;Cross-references: UNIPROT:P43667; UNIPARC:UPI000004F5DE; GB:AE000353; GB:U000096; NID

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: ysaH

C;Superfamily: Escherichia coli hypothetical protein (emrR 5' region)

Query Match	100.0%	Score 552;	DB 2;	Length 111;
Best Local Similarity	100.0%;	Pred. No. 1.5e-49;		
Matches 111;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSYEVLGLLVGVNYCFRYPFLRLRVGNARPTKRGAVGILLDTIGIASICALLVNSTA	60	
Db	1	MSYEVLGLLVGVNYCFRYPFLRLRVGNARPTKRGAVGILLDTIGIASICALLVNSTA	60	
Qy	61	PEVMDTRRFPVTLVGFVAVLGASFYKTRSIIPITLLSALAYGLAWKVMII	111	
Db	61	PEVMDTRRFPVTLVGFVAVLGASFYKTRSIIPITLLSALAYGLAWKVMII	111	

RESULT 2

A91072

hypothetical protein ECs3545 [imported] - Escherichia coli (strain O157:H7, substrain R

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: A91072

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gaasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: A91072

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-111 <HAY>

A;Cross-references: UNIPROT:Q8X907; UNIPARC:UPI000000E4F0; GB:BA000007;

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A:Gene: ECa3545
C:Superfamily: Escherichia coli hypothetical protein (emrR 5' region)

Query Match 99.3%; Score 548; DB 2; Length 111;
Best Local Similarity 99.1%; Pred. No. 3.9e-49;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSYEVLLGLLVGVNYCFRVLPLRLRVGNARPTKRGAVGILLDTIGTIGASTICALLVSTA 60
Db 1 MSYEVLLGLLVGVNYCFRVLPLRLRVGNARPTKRGAVGILLDTIGTIGASTICALLVSTA 60

Qy 61 PEVMDHTRFRFVPTLVGFVAVLGASFYKTRSIITPTLLSALAYGLAWKWMII 111
Db 61 PEVMDHTRFRFVPTLVGFVAVLGASFYKTRSIITPTLLSALAYGLAWKWMII 111

RESULT 3
D85916
hypothetical protein ygaH [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85916
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85916
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-111 <KUR>
A:Cross-references: UNIPROT:Q8X907; UNIPARC:UPI0000165BDF; GB:AE005174; NID:gl2517124; P
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ygaH
C:Superfamily: Escherichia coli hypothetical protein (emrR 5' region)

Query Match 98.6%; Score 544; DB 2; Length 111;
Best Local Similarity 98.2%; Pred. No. 1e-48;
Matches 109; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSYEVLLGLLVGVNYCFRVLPLRLRVGNARPTKRGAVGILLDTIGTIGASTICALLVSTA 60
Db 1 MSYEVLLGLLVGVNYCFRVLPLRLRVGNARPTKRGAVGILLDTIGTIGASTICALLVSTA 60

Qy 61 PEVMDHTRFRFVPTLVGFVAVLGASFYKTRSIITPTLLSALAYGLAWKWMII 111
Db 61 PEVMDHTRFRFVPTLVGFVAVLGASFYKTRSIITPTLLSALAYGLAWKWMII 111

RESULT 4
AG0396
probable membrane protein YPO3265 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG0396
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraghi, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0396
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <KUR>
A:Cross-references: UNIPROT:Q8ZBW9; UNIPARC:UPI00000DC73C; GB:AL590842; PIDN:CAC92499.1;
C:Genetics:
A:Gene: YPO3265
C:Superfamily: Escherichia coli hypothetical protein (emrR 5' region)

Query Match 55.9%; Score 308.5; DB 2; Length 113;
Best Local Similarity 54.4%; Pred. No. 1.4e-24;

Matches 62; Conservative 21; Mismatches 26; Indels 5; Gaps 2;

Qy 1 MSYEVLLGLLVGVNYCFRVLPLRLRVGNARPTK---RGAVGILLDTIGTIGASTICALLV 57
Db 1 MNTDILMIGLVGVNYCFRVLPLRLRVGNARPTK---GPAKQTGLLRGMKSLDLSIGIASICALLV 58

Qy 58 STAPEVMDHTRFRFVPTLVGFVAVLGASFYKTRSIITPTLLSALAYGLAWKWMII 111
Db 59 SSTPEIHHFQKLLPTLIGLVICGCFYKTRSIITPATLLGALCYGLTFKLLTIL 112

RESULT 5
A95272
hypothetical protein Sma0151 [imported] - Sinorhizobium meliloti (strain 1021) magaplasma
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: A95272
R:Barnett, M.J.; Fisher, R.P.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.: Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: A95272
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-866 <KUR>
A:Cross-references: UNIPROT:Q930W3; UNIPARC:UPI000000CAF89; GB:AE006469; PIDN:AAK64739.1;
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.P.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heBault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.;
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma0151
A:Genome: plasmid

Query Match 14.2%; Score 78.5; DB 2; Length 866;
Best Local Similarity 25.4%; Pred. No. 3.7;
Matches 30; Conservative 17; Mismatches 40; Indels 31; Gaps 4;

Qy 13 GVANYCFRVLPLRLRVGNA-----RPTKRGAVGILLDTIGTIGASI----- 51
Db 4 GIANLAYVSIPIFMVLMGAAVASSPAGSDLYTSLDRWLNRIFGGLILSNIGACAIIFSGMTG 63

Qy 52 -----CALLVVSTAPEVMDHTRFRFVPTLVGFVAVLGASFYKTRSIITPTLLSALAYGLA 104
Db 64 SSPATCAIGKMGIPENM---RRGYPASVASGSIAG--GTLGILIPPSVTLIVYGLA 116

RESULT 6
G87022
hypothetical protein murE [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004
C:Accession: G87022
R:Cole, S.T.; Biglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho;
am, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: G87022
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-530 <STO>
A:Cross-references: UNIPROT:O69557; UNIPARC:UPI000012F9A0; GB:AL450380; NID:gl13092979; P
C:Genetics:

C;Genetics:
A;Gene: lin0435
C;Superfamily: gufA protein

Query Match 13.2%; Score 73; DB 2; Length 269;
Best Local Similarity 33.7%; Pred. No. 4.3;
Matches 34; Conservative 13; Mismatches 34; Indels 20; Gaps 6;
QY 6 LLLGLLVGV-----ANYCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLYV 56
Db 11 VLLALLAGIPTWACTAAGASIVFFFKLNKKWGNV--MLGFAAGVWL-----AASFWSLL- 63
QY 57 VSTAPEVNHDTTR--FVPTLVGPAVLGASFYKTRSIITPL 95
Db 64 -APAIEMSKOLGKFSFVPALWGF-LLGGIFLRVIDRIIPL 102

RESULT 15

S28025
light harvesting complex II assembly factor pucC [imported] - Rhodobacter sphaeroides
C;Species: Rhodobacter sphaeroides
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: S28025; T50701
R;Gibson, L.C.D.; McGlynn, P.; Chaudhri, M.; Hunter, C.N.
Mol. Microbiol. 6, 3171-3186, 1992
A;Title: A putative anaerobic coproporphyrinogen III oxidase in Rhodobacter sphaeroides.
A;Reference number: S28023; MUID:93086425; PMID:1453956
A;Accession: S28025
A;Molecule type: DNA
A;Residues: 1-459 <GIB>
A;Cross-references: UNIPROT:Q02443; UNIPARC:UPI00001329F0; EMBL:X68796; NID:G46438; PIDN
R;Choudhary, M.; Kaplan, S.
Nucleic Acids Res. 28, 862-867, 2000
A;Title: DNA sequence analysis of the photosynthesis region of Rhodobacter sphaeroides 2
A;Reference number: Z25222; MUID:20115911; PMID:10648776
A;Accession: T50701
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-459 <CHO>
A;Cross-references: UNIPARC:UPI00001329F0; EMBL:AF195122; PIDN:AAF24245.1
A;Experimental source: strain 2.4.1
C;Genetics:
A;Gene: pucC

Query Match 13.2%; Score 73; DB 2; Length 459;
Best Local Similarity 29.0%; Pred. No. 7.3;
Matches 31; Conservative 11; Mismatches 51; Indels 14; Gaps 3;
QY 5 VLLLG-LLVGNAVYCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLYVSTAPEV 63
Db 352 VFLIGTLVVGFGGLFSGHTL---TATWRLAPKEQVGLALGAWGAVQATAAGVAIAGAV 408
QY 64 MHDTRRFVPTLVG-----PAVLGASFYKTRSIITPLSALA 100
Db 409 LRDILOAMPDLGSGYGPAGYVAVFAEGFLFTMTIVILPLRSALA 455

Search completed: February 15, 2006, 12:24:27
Job time : 16.9663 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 12:16:07 ; Search time 91.9803 Seconds
(without alignments)
851.417 Million cell updates/sec

Title: US-10-073-293A-6
Perfect score: 552
Sequence: 1 MSFEVLLGLLVGVNYCFR.....IPTLLSALAYGLAWKVMAII 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	552	100.0	111	1 YGAH_ECOLI	P43667 escherichia
2	548	99.3	111	2 Q7ABD5_ECO57	Q7abds escherichia
3	548	99.3	111	2 Q8FEQ5_ECOL6	Q8fed5 escherichia
4	548	99.3	111	2 Q83JZ5_SHIFL	Q83jz5 shigella fl
5	544	98.6	111	2 Q8X907_ECO57	Q8x907 escherichia
6	313.5	56.8	113	2 Q93KB3_ERWCH	Q93kb3 erwinia chr
7	309.5	56.1	113	2 Q66E35_YERPS	Q66e35 yerinia ps
8	309.5	56.1	113	2 Q6D1D8_ERWCT	Q6dlld8 erwinia car
9	308.5	55.9	113	2 Q8ZBW9_YERPE	Q8zbw9 yerinia pe
10	271.5	49.2	112	2 Q7N780_PHOLL	Q7n780 photorhabdu
11	87	15.8	408	2 Q6MH96_EDEBA	Q6mh96 bdellovibri
12	86.5	15.7	983	2 Q4NUH1_DDELT	Q4nuhi anaeromyxob
13	86	15.6	463	2 Q9HIG1_THEAC	Q9hig1 thermoplasma
14	86	15.6	478	2 Q4V7J8_XENLA	Q4v7j8 xenopus lae
15	83.5	15.1	169	2 Q5TAJ8_HUMAN	Q5taj8 homo sapien
16	82.5	14.9	167	2 Q9N3B7_CAELI	Q9n3b7 caenorhabdi
17	82	14.9	575	2 Q98S82_RHILO	Q98s82 rhizobium l
18	82	14.9	587	2 Q7UQM5_RHOBA	Q7uqms rhodospirell
19	82	14.9	589	1 FBX24_MOUSE	Q9d417 mus musculu
20	81.5	14.8	297	2 Q982L1_RHILO	Q982l1 rhizobium l
21	80.5	14.6	400	1 GTR3_RABIT	Q9xsc2 oryctolagus
22	79.5	14.4	355	2 Q8TH78_METAC	Q8th78 methanoearc
23	79	14.3	366	2 Q83G83_TROWT	Q83g83 tropheryma
24	79	14.3	366	2 Q83HY7_TROW8	Q83hy7 tropheryma
25	78.5	14.2	238	2 Q61B17_CAEBR	Q61b17 caenorhabdi
26	78.5	14.2	495	1 GTR3_CANFA	P47842 canis famil
27	78.5	14.2	516	2 Q8EH66_SHEON	Q8eh66 shewanella
28	78.5	14.2	611	2 Q4HVS1_GIBZE	Q4hvs1 gibberella
29	78.5	14.2	866	2 Q930W3_RHIME	Q930w3 rhizobium m
30	78	14.1	530	1 MURE_MYCLE	Q69557 mycobacteri
31	78	14.1	538	2 Q5YTQ8_NOCPA	Q5ytq8 nocardia fa

32	77.5	14.0	365	2	Q872P7_NEUCR	Q872p7 neurospora
33	77.5	14.0	437	2	Q4KFY4_PSEF5	Q4kfy4 pseudomonas
34	77.5	14.0	496	1	GTR3_PONPY	Q5r608 pongo pygma
35	77	13.9	138	2	Q72CW5_DESVH	Q72cw5 desulfovibr
36	77	13.9	358	2	Q7SA63_NEUCR	Q7sa63 neurospora
37	77	13.9	535	2	Q8RVG5_DAUCA	Q8rvgs daucus caro
38	76.5	13.9	309	1	YQGH_BACSU	P46339 bacillus su
39	76.5	13.9	433	2	Q5AWM6_EMENI	Q5awm6 aspergillus
40	76.5	13.9	494	1	GTR3_BOVIN	P58352 bos taurus
41	76.5	13.9	494	1	GTR3_SHEEP	P47843 ovis aries
42	76	13.8	475	2	Q8ABF2_BACTN	Q8abf2 bacteroides
43	76	13.8	502	2	Q93Z26_ARATH	Q93z26 arabidopsis
44	76	13.8	528	2	Q4UW01_XANCP	Q4uw01 xanthomonas
45	76	13.8	528	2	Q8P842_XANCP	Q8p842 xanthomonas

ALIGNMENTS

RESULT 1

ID	YGAH_ECOLI	STANDARD;	PRT;	111 AA.
AC	P43667; P77323;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Hypothetical protein YGAH.			
GN	NamesygaH; OrderedLocuNames=b2683;			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
OX	NCBI_TaxID=562;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=K12 / MG1655;			
RX	MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.";			
RL	Science 277:1453-1474 (1997).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=K12;			
RX	MEDLINE=97349980; PubMed=9205837;			
RA	Yamanoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,			
RA	Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,			
RA	Mizobuchi K., Mori H., Nakade S., Nakamura Y., Washimoto H.,			
RA	Oshina T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,			
RA	Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,			
RA	Yamagata S., Horiuchi T.;			
RT	"Construction of a contiguous 874-kb sequence of the Escherichia coli-			
RT	K12 genome corresponding to 50.0-68.8 min on the linkage map and			
RT	analysis of its sequence features.";			
RT	DNA Res. 4:91-113 (1997).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE OF 39-111.			
RC	STRAIN=K12;			
RX	MEDLINE=95247664; PubMed=7730261;			
RA	Lomovskaya O., Lewis K., Matin A.;			
RT	"EmrR is a negative regulator of the Escherichia coli multidrug			
RT	resistance pump EmrAB.";			
RL	J. Bacteriol. 177:2328-2334 (1995).			
RN	[4]			
IDENTIFICATION.				
RA	Rudd K.E.;			
RA	Unpublished observations (JUL-1995).			
CC	-----			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use as long as its content is in no way modified and this statement is			

```

CC removed.
CC -----
DR EMBL; U00096; AAC75730.1; -; Genomic_DNA.
DR EMBL; D90891; BAA16545.1; -; Genomic_DNA.
DR EMBL; U19993; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR PIR; D65048; D65048. -;
DR ECHOBASE; EB2775; -;
DR EcoGene; EG12940; yghH.
KW Complete proteome; yghH.
FT CONFLICT 48 49 IA -> MP (in Ref. 3).
FT CONFLICT 54 54 L -> Q (in Ref. 3).
FT CONFLICT 89 111 SIIPTLLSALAYGLAWKVMII -> THYHPNTA (in Ref. 3).
FT SEQUENCE 111 AA; 12024 MW; 2527D05A3E1DA69D CRC64;

Query Match 100.0%; Score 552; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.5e-44;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSYEVLLGLLVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVWSTA 60
Db 1 MSYEVLLGLLVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVWSTA 60

Qy 61 PEVMDTRFRVPTLVGPAVLGASFYKTRSIIPILLSALAYGLAWKVMII 111
Db 61 PEVMDTRFRVPTLVGPAVLGASFYKTRSIIPILLSALAYGLAWKVMII 111

RESULT 2
Q7ABD5_ECO57 PRELIMINARY; PRT; 111 AA.
ID Q7ABD5;
AC Q7ABD5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein ECG3545.
GN OrderedLocusNames=ECG3545;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba H., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL; BA000007; BAB36968.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 111 AA; 11995 MW; 3E392E54DEF3A863 CRC64;

Query Match 99.3%; Score 548; DB 2; Length 111;
Best Local Similarity 99.1%; Pred. No. 3.7e-44;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSYEVLLGLLVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVWSTA 60
Db 1 MSYEVLLGLLVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVWSTA 60

Qy 61 PEVMDTRFRVPTLVGPAVLGASFYKTRSIIPILLSALAYGLAWKVMII 111
Db 61 PEVMDTRFRVPTLVGPAVLGASFYKTRSIIPILLSALAYGLAWKVMII 111

RESULT 3
Q8FEQ5_ECOL6 PRELIMINARY; PRT; 111 AA.
ID Q8FEQ5;
AC Q8FEQ5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein yghH.
GN Name=yghH; OrderedLocusNames=c3236;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Rosch P.,
RA Raiko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Ferna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016765; AA81688.1; -; Genomic_DNA.
DR PIR; A91072; A91072.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 111 AA; 11995 MW; 3E392E54DEF3A863 CRC64;

Query Match 99.3%; Score 548; DB 2; Length 111;
Best Local Similarity 99.1%; Pred. No. 3.7e-44;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSYEVLLGLLVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVWSTA 60
Db 1 MSYEVLLGLLVGVANVCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVWSTA 60

Qy 61 PEVMDTRFRVPTLVGPAVLGASFYKTRSIIPILLSALAYGLAWKVMII 111
Db 61 PEVMDTRFRVPTLVGPAVLGASFYKTRSIIPILLSALAYGLAWKVMII 111

RESULT 4
Q83JZ5_SHIFL PRELIMINARY; PRT; 111 AA.
ID Q83JZ5;
AC Q83JZ5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein yghH.
GN Name=yghH; OrderedLocusNames=S2897, SF2710;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Ferna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella

```

```

RT flexneri serotype 2a strain 2457T. ";
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL; AE005674; AAN44203.1; -; Genomic DNA.
DR EMBL; AE016987; AAP18030.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 111 AA; 11995 MW; 38392E54DEF3A863 CRC64;

Query Match
Best Local Similarity 99.3%; Score 548; DB 2; Length 111;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSYEVLGGLGVGNYCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVYSTA 60
Db 1 MSYEVLGGLGVGNYCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVYSTA 60
Qy 61 PEVNHDTTRFRVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKVMII 111
Db 61 PEVNHDTTRFRVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKVMII 111

RESULT 5
Q8X907 ECO57
ID Q8X907 ECO57 PRELIMINARY; PRT; 111 AA.
AC Q8X907;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein ygaH.
GN Name=ygaH; OrderedLocusNames=z3984;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHRC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
DR EMBL; AE005174; AAG57792.1; -; Genomic DNA.
DR PIR; A91072; A91072.
DR PIR; D85916; D85916.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 111 AA; 12018 MW; 3E392E54DEF3A184 CRC64;

Query Match
Best Local Similarity 98.6%; Score 544; DB 2; Length 111;
Matches 109; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSYEVLGGLGVGNYCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVYSTA 60
Db 1 MSYEVLGGLGVGNYCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVYSTA 60
Qy 61 PEVNHDTTRFRVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKVMII 111
Db 61 PEVNHDTTRFRVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKVMII 111

RESULT 6
Q93KB3 ERWCH
ID Q93KB3 ERWCH PRELIMINARY; PRT; 113 AA.
AC Q93KB3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE YgaH protein.
GN Name=ygaH;

```

```

OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3937;
RA Reverchon S.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ410307; CAC43438.1; -; Genomic DNA.
SQ SEQUENCE 113 AA; 12338 MW; CC18A5FF9C9523D CRC64;

Query Match
Best Local Similarity 58.9%; Score 313.5; DB 2; Length 113;
Matches 63; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

Qy 1 MSYEVLGGLGVGNYCFRYLPLRLRVGNARPTKRGAVGILLDTIGIASICALLVYST 59
Db 1 MNTSVLLIGLVGTVNPLFRYLPLRLGAGRTSGNLRGRGTALLDSDIGIASICALLVYST 60
Qy 60 APEVMDHTRFRVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWK 106
Db 61 LPDIMQHTKXMLPTLAGFALLTLCFYKTRSIIVLSTLLGALCYGIVFK 107

RESULT 7
Q66E35 YERPS
ID Q66E35 YERPS PRELIMINARY; PRT; 113 AA.
AC Q66E35;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative LIV-E family branched chain amino acid exporter, small
DE subunit.
GN Name=ygaH; OrderedLocusNames=YPTB0858;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831 (2004).
DR EMBL; BX936398; CAH20098.1; -; Genomic DNA.
KW Complete proteome.
SQ SEQUENCE 113 AA; 12300 MW; F244135984358C8A CRC64;

Query Match
Best Local Similarity 56.1%; Score 309.5; DB 2; Length 113;
Matches 63; Conservative 20; Mismatches 26; Indels 5; Gaps 2;

Qy 1 MSYEVLGGLGVGNYCFRYLPLRLRVGNARPTK---RGAVGILLDTIGIASICALLV 57
Db 1 MNTDVLMIGLVGVTVNPLFRYLPLRLGAGRTSGNLRGRGTALLDSDIGIASICALLV 58
Qy 58 STAPEVMDHTRFRVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKVMII 111
Db 59 SSTPEIIHPQKLLPTLIGLVICGCFYKNTSIIFATLLGALCYGTFKLLTIL 112

RESULT 8
Q6D1D8 ERWCT
ID Q6D1D8 ERWCT PRELIMINARY; PRT; 113 AA.
AC Q6D1D8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)

```

```
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=EC3510;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holleva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RA "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
DR EMBL; BX950851; CAG76408.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 113 AA; 12232 MW; EAA747D0E5BA18E7 CRC64;

Query Match 56.1%; Score 309.5; DB 2; Length 113;
Best Local Similarity 56.0%; Pred. No. 1.5e-21;
Matches 61; Conservative 21; Mismatches 26; Indels 1; Gaps 1;

Qy 1 MSYEVLLGLLVGVNYCFRYLPLRLRVGNARPT--KRGAVGILLDTTIGIASICALLVYST 59
Db 1 MSTEVLLGLLVGVNYCFRYLPLRLRVGNARPT--KRGAVGILLDTTIGIASICALLVYST 60

Qy 60 APEVMHDTRRFVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKVM 108
Db 61 VPDILTTHKLPITLVGPIITLTACFYKTRSIIVLSTLLGAFYCGIAFKLL 109

RESULT 9
ID Q8ZBW9 YERPE PRELIMINARY; PRT; 113 AA.
AC Q8ZBW9 Q74X00; Q7CK79;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative membrane protein (Hypothetical protein y0924).
GN OrderedLocusNames=YPO667, YPO3265, Y0924;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1126/SCIENCE.1184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
```

```
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans.";
RL DNA Res. 11:179-197(2004).
DR EMBL; AJ414156; CAC92499.1; -; Genomic_DNA.
DR EMBL; AE013695; AAM84506.1; -; Genomic_DNA.
DR EMBL; AE017129; AAS60934.1; -; Genomic_DNA.
DR PIR; AG0396; AG0396.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 113 AA; 12314 MW; 1FF2DF284D0406A2 CRC64;

Query Match 55.9%; Score 308.5; DB 2; Length 113;
Best Local Similarity 54.4%; Pred. No. 1.9e-21;
Matches 62; Conservative 21; Mismatches 26; Indels 5; Gaps 2;

Qy 1 MSYEVLLGLLVGVNYCFRYLPLRLRVGNARPTK---RGAAGVILLDTTIGIASICALLV 57
Db 1 MNTDILMIGLVGVNYVFLRYLPLRF--GPARKQTGLLRGKMSLLDSIGIASICALLV 58

Qy 58 STAPEVMHDTRRFVPTLVGFAVLGASFYKTRSIIPITLLSALAYGLAWKVM 111
Db 59 SSFPEIHHKLPITLVGPIITLTACFYKTRSIIVLSTLLGALCYGLTFKLTIL 112

RESULT 10
ID Q7N780 PHOLL PRELIMINARY; PRT; 112 AA.
AC Q7N780;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to unknown protein YgaH of Escherichia coli.
GN OrderedLocusNames=plu1278;
OS Photobacterium luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photobacterium.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taurit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freysinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Signier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boenare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photobacterium
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
DR EMBL; BX571863; CAB13572.1; -; Genomic_DNA.
KW Photolyst; plu1278; -.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12245 MW; 570F8F8E73247F8 CRC64;

Query Match 49.2%; Score 271.5; DB 2; Length 112;
Best Local Similarity 50.5%; Pred. No. 5.9e-18;
Matches 54; Conservative 23; Mismatches 25; Indels 5; Gaps 2;

Qy 4 EVLLGLLVGVNYCFRYLPLRLRVGNARPT---KRGAVGILLDTTIGIASICALLVSTA 60
Db 5 KILLIGLVGLNPFYRLPLRF--GKAROSAGRKAGKTSIILDSIGIASICSLIVSGV 62
```

Qy 61 PEVMDTRFRVPTLVGPAVLGASFYKTRSIIPITLLSALAYGLAWKV 107
 Db 63 PDVMSQKLLPTLIGCLTICLVFYKTRKQIIILATLFGALFLGFLFKI 109

RESULT 11
 Q6MH96_BDEBA PRELIMINARY; PRT; 408 AA.
 AC Q6MH96;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE MFS permease.
 GN Name=ydr; OrderedLocusNames=Bd3656;
 OS Bdellovibrio bacteriovorus.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
 OC Bdellovibrionaceae; Bdellovibrio.
 OX NCBI_TaxID=959;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
 RX PubMed=14752164; DOI=10.1126/science.1093027;
 RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
 RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,
 RA Sockett R.E., Schuster S.C.;
 RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
 RT genomic perspective";
 RL Science 303:689-692(2004).
 DR EMBL; BX842656; CAE81031.1; -; Genomic_DNA.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR011701; MFS_1.
 DR InterPro; IPR000209; Pfam_S8_S53.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN 1.
 KW Complete proteome.
 SQ SEQUENCE 408 AA; 43528 MW; A353E3898A344610 CRC64;

Query Match 15.8%; Score 87; DB 2; Length 408;
 Best Local Similarity 28.4%; Pred. No. 5.6;
 Matches 31; Conservative 20; Mismatches 54; Indels 4; Gaps 3;

Qy 5 VLLGLGAVG--VANYCFRYPLRLRVGNARPTKRGAVGILLDTIGIASICALLVWSTAPE 62
 Db 27 VLLMAAVGIIVANLYAQ-PITAMISQALGLDPSAAGLVVTLTQIGYGLGVLLIVPLGD 85

Qy 63 VMHDTFRFVPTLVGPAVLGASFYKTRSIIPITLLSALAYGLAWKVMAII 111
 Db 86 II-ENRLVLTMIAGVLGVLGAFASQLTPFYFAATGATGLGASTVQIL 133

RESULT 12
 Q4NUH1_9DELT
 ID Q4NUH1_9DELT PRELIMINARY; PRT; 983 AA.
 AC Q4NUH1;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Spermine synthase precursor.
 GN ORFNames=AdelhdRAFT_2712;
 OS Anaeromyxobacter dehalogenans 2CP-C.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 OC Cytophactereae; Myxococcaceae; Anaeromyxobacter.
 OX NCBI_TaxID=290397;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=2CP-C;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,

RA Hamon N., Israni S., Pitluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of Anaeromyxobacter
 RT dehalogenans 2CP-C.";
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=2CP-C;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome assembly of Anaeromyxobacter
 RT dehalogenans 2CP-C."; to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AHD01000014; EAL79250.1; -; Genomic_DNA.
 KW Signal.
 FT SIGNAL
 SQ SEQUENCE 983 AA; 105076 MW; 66DF12DF091816DA CRC64;

Query Match 15.7%; Score 86.5; DB 2; Length 983;
 Best Local Similarity 27.0%; Pred. No. 13;
 Matches 30; Conservative 22; Mismatches 52; Indels 7; Gaps 4;

Qy 2 SYEVLGLGLLVGVANYCFRYPLRLRVGNARPTKRGAVGILLDTIGIASICALLVWSTAP 61
 Db 269 SFTVMLVAFLVGLGALGSYAFGVPRLFGGERHAGVLVFGCIQVLIGVSALAVTLMRDLP 328

Qy 62 EVMDHTRFRVPT--TLVGFAV-LGASP-YKTRSIIPITLLSALAY---GLAW 105
 Db 329 LTAGHTRALLAGLELTRFAERVSGSFGIAFTYLLIPSLMGLAPFVAGAVW 379

RESULT 13
 Q9HGL1_THEAC
 ID Q9HGL1_THEAC PRELIMINARY; PRT; 463 AA.
 AC Q9HGL1;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein tai1378.
 GN OrderedLocusNames=Tai1378;
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.
 OX NCBI_TaxID=2303;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001; DOI=10.1038/35035069;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
 RT acidophilum.";
 RL Nature 407:508-513(2000).
 DR EMBL; AL445067; CAC12499.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR002293; AA/rel_permease1.
 DR InterPro; IPR004841; Permease_region.
 DR Pfam; PF00324; AA_permease; 1.
 KW Complete proteome; Hypothetical protein; Transmembrane; Transport.
 SQ SEQUENCE 463 AA; 48756 MW; F6DFBE0B3C619DC1 CRC64;

Query Match 15.6%; Score 86; DB 2; Length 463;
 Best Local Similarity 33.3%; Pred. No. 7.8;
 Matches 24; Conservative 17; Mismatches 25; Indels 6; Gaps 3;

Qy 45 TIGIASICALLVWSTAPEVMDHTRFRF---VPTLVGFAVLGASFYKT-RSIIPTLLSAL 99
 Db 45 TIGIASICALLVWSTAPEVMDHTRFRF---VPTLVGFAVLGASFYKT-RSIIPTLLSAL 99


```

; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: MelkieJohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-299-549-5

      Query Match          13.9%; Score 76.5; DB 2; Length
      Best Local Similarity 34.1%; Pred. No. 1.1;
      Matches 29; Conservative 17; Mismatches 30; Indel

QY      2 SYEVLILG-LLVGV-ANYCFRYPILRLRVGNARPTK-RGAVGILLD
Db      116 SVEMLIIGRLIIGLCGLCTGFVP--WYIGISPTALRGAPG-TLN-
QY      59 T-APEVMHDTTRFVPTLVGFVILGA 82
Db      171 IFGLKVLIGTDLWPLLIGFTILPA 195

RESULT 5
US-09-610-417-5
; Sequence 5, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95

```


;; SOFTWARE: FastSeq for Windows Version 2.0
;; APPLICATION DATA:
;; FILING DATE: 05-Jul-2000
;; PRIOR APPLICATION DATA:
;; FILING DATE: 09/299,549
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meikiejohn, Ph.D., Anita L.
;; REGISTRATION NUMBER: 35,283
;; REFERENCE/DOCKET NUMBER: 07334/072002
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617/542-5070
;; TELEFAX: 617/542-8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 494 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-610-417-5

Query Match 13.9%; Score 76.5; DB 2; Length 494;
Best Local Similarity 34.1%; Pred. No. 1.1;
Matches 29; Conservative 17; Mismatches 30; Indels 9; Gaps 7;

QY 2 SYEVLILG-LLVGV-ANYCFRYLPLRLVGNARPTK-RGAVGILLDTGTIGIASICALVVS 58
Db 116 SVEMLILGRLLIGLFCGLCTGFVP--MYIGISPTALRGAGF-TLNQLGI--VIGILVAQ 170
QY 59 T-APEVMHDTRRFVPTLVGFAVLGA 82
Db 171 IFGLKVILGTEDLWPLLGLFTILPA 195

RESULT 6
US-09-981-947B-5
; Sequence 5, Application US/09981947B
; Patent No. 6933115
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding GLUTX
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: MPI1998-021DV3
; CURRENT APPLICATION NUMBER: US/09/981,947B
; CURRENT FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 09/610,417
; PRIOR FILING DATE: 2000-07-05
; PRIOR APPLICATION NUMBER: 09/299,549
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 09/031,392
; PRIOR FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 494
; TYPE: PRT
; ORGANISM: human
US-09-981-947B-5

Query Match 13.9%; Score 76.5; DB 2; Length 494;
Best Local Similarity 34.1%; Pred. No. 1.1;
Matches 29; Conservative 17; Mismatches 30; Indels 9; Gaps 7;

QY 2 SYEVLILG-LLVGV-ANYCFRYLPLRLVGNARPTK-RGAVGILLDTGTIGIASICALVVS 58
Db 116 SVEMLILGRLLIGLFCGLCTGFVP--MYIGISPTALRGAGF-TLNQLGI--VIGILVAQ 170
QY 59 T-APEVMHDTRRFVPTLVGFAVLGA 82
Db 171 IFGLKVILGTEDLWPLLGLFTILPA 195

Db 171 IFGLKVILGTEDLWPLLGLFTILPA 195

RESULT 7
US-09-134-001C-5531
; Sequence 5531, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5531
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5531

Query Match 13.8%; Score 76; DB 2; Length 655;
Best Local Similarity 21.0%; Pred. No. 1.8;
Matches 34; Conservative 21; Mismatches 39; Indels 68; Gaps 7;

QY 5 VLLGLLVGV-----ANYCFRYLPLRLVGNARPTKRGAVGILLDTGTIGIASIC 52
Db 479 IMILGLVIGANMAIDMGPFNKAAVYF--ATAALVTGNAAAPITAAMIGMIPPLAIAT-- 534
QY 53 ALLVW-----STAPEVMHDTRRFVPT-LVG----- 76
Db 535 AMLIFRRKTYKQSGSIVPNVNGLSFITEGALPFFAAADPLRVIPSMVGVGAGATLG 594
QY 77 -----FAVLGASF-----YKTRSIITLLSALAYG 102
Db 595 LGSSIRKAPHGGIFVIIGTDFNHLIQTLLALVVGTLVSALIYG 636

RESULT 8
US-09-603-208A-228
; Sequence 228, Application US/09603208A
; Patent No. 6822084
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; FILE REFERENCE: BGI-124CP
; CURRENT APPLICATION NUMBER: US/09/603,208A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08


```

Qy      49   ---ASICALLVVSTAPEVMHTRRPVPVTLV--GFAVLGASPYKTRSIILIPTLLSALA-YG 102
         ||| : ||| : |:: :|| : ||| : ||| : ||| :
Db     143   GTAANMAGLLPSADNPV---DSQTLLIISMVTGLGVTLGSVMFRGFALIPILLIGVLGYA 199
         ||| : ||| : |:: :|| : ||| : ||| : ||| :
Qy     103   LAW 105
         ||| :
Db     200   LSF 202
         ||| :

RESULT 14
US-09-605-703B-1560
; Sequence 1560, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1560
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-1560
```

[illegible]

```
Db      74 SAKKIVLPVAVGVTVGTASINADAAFTVVVPPLAAMVFKAAGRHPVAGLLIGSFAAVGAGY- 130
Qy      86 KTRSIIPTTLISALAYGLAWKM 108
        :|||:|||:
Db     133 --STAIVPTSILDALFAGITNAVM 153

RESULT 15
US-09-328-352-7553
; Sequence 7553, Application US/09328352
; Patent NO. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7553
; LENGTH: 294
; TYPE: PRP
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7553
```

```

; SEQ ID NO 7553
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7553

Query Match      12.7%; Score 70; DB 2; Length 294;
Best Local Similarity 25.8%; Pred. No. 3.5;
Matches 33; Conservative 19; Mismatches 48; Indels 28; Gaps 5

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 12:24:42 ; Search time 71.4017 Seconds
(without alignments)
649.551 Million cell updates/sec

Title: US-10-073-293A-6

Perfect score: 552

Sequence: 1 MSYEVLLGLLVGVNYCFR.....IPTLLSALAYGLAWKVMII 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA Main:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
 - 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
 - 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
 - 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	552	100.0	111	5	US-10-073-293A-6
2	84	15.2	223	4	US-10-424-599-223037
3	80.5	14.6	400	5	US-10-732-923-23642
4	79.5	14.4	1238	4	US-10-424-599-161652
5	78.5	14.2	495	5	US-10-732-923-23556
6	78	14.1	530	4	US-10-282-122A-63783
7	77.5	14.0	442	4	US-10-282-122A-62944
8	77	13.9	308	4	US-10-369-493-3529
9	76.5	13.9	161	5	US-10-732-923-23831
10	76.5	13.9	494	3	US-09-981-947A-5
11	76.5	13.9	494	5	US-10-732-923-23560
12	76.5	13.9	494	5	US-10-732-923-23639
13	76.5	13.9	500	4	US-10-369-493-12445
14	76	13.8	509	4	US-10-092-900A-314
15	76	13.8	538	4	US-10-437-963-107875
16	76	13.8	650	4	US-10-282-122A-70668
17	76	13.8	655	4	US-10-724-972A-7255
18	75.5	13.7	201	5	US-10-501-282-5130
19	75.5	13.7	225	5	US-10-501-282-5132
20	75.5	13.7	253	5	US-10-501-282-5134
21	75.5	13.7	254	5	US-10-501-282-5136
22	75.5	13.7	283	5	US-10-501-282-5138
23	75.5	13.7	326	5	US-10-501-282-5140
24	75.5	13.7	327	5	US-10-501-282-5142
25	75.5	13.7	362	4	US-10-437-963-115286
26	75.5	13.7	497	5	US-10-732-923-23714
27	75.5	13.7	520	5	US-10-732-923-23713

28	74.5	13.5	551	5	US-10-968-848-84	Sequence 84, Appl
29	74	13.4	250	4	US-10-425-115-215825	Sequence 215825,
30	74	13.4	472	4	US-10-080-170-380	Sequence 380, App
31	74	13.4	472	4	US-10-080-170-380	Sequence 380, App
32	74	13.4	472	4	US-10-468-356-380	Sequence 380, App
33	74	13.4	791	3	US-09-738-626-4530	Sequence 4530, Ap
34	73.5	13.3	156	5	US-10-852-335A-111	Sequence 111, App
35	72.5	13.1	496	4	US-10-170-385-247	Sequence 247, App
36	72.5	13.1	496	4	US-10-341-434-89	Sequence 89, Appl
37	72.5	13.1	496	4	US-10-755-889-666	Sequence 666, App
38	72.5	13.1	496	5	US-10-893-315-85	Sequence 85, Appl
39	72.5	13.1	499	5	US-10-893-315-92	Sequence 92, Appl
40	72	13.0	185	4	US-10-425-115-342079	Sequence 342079,
41	72	13.0	481	4	US-10-437-963-199115	Sequence 199115,
42	71.5	13.0	225	4	US-10-282-122A-53684	Sequence 53684, A
43	71.5	13.0	488	5	US-10-732-923-11192	Sequence 11192, A
44	71.5	13.0	489	5	US-10-732-923-23545	Sequence 23545, A
45	71.5	13.0	502	4	US-10-156-761-12709	Sequence 12709, A

ALIGNMENTS

RESULT 1

US-10-073-293A-6

; Sequence 6, Application US/10073293A

; Publication No. US20050239175A1

; GENERAL INFORMATION:

; APPLICANT: TABOLINA, EKATERINA

; APPLICANT: RYBAK, KONSTANTIN

; APPLICANT: KHOURGES, EVGENI

; APPLICANT: VOROSHILOVA, ELVIRA

; APPLICANT: GUSYATINER, MIKHAIL

; TITLE OF INVENTION: METHOD FOR PRODUCING L-AMINO ACID USING BACTERIA BELONGING TO TI

; TITLE OF INVENTION: ESCHERICHIA

; FILE REFERENCE: 219594US0

; CURRENT APPLICATION NUMBER: US/10/073,293A

; CURRENT FILING DATE: 2002-02-13

; PRIOR APPLICATION NUMBER: RU 2001103865

; PRIOR FILING DATE: 2001-02-13

; PRIOR APPLICATION NUMBER: RU 2001104998

; PRIOR FILING DATE: 2001-02-26

; PRIOR APPLICATION NUMBER: RU 2001104999

; PRIOR FILING DATE: 2001-02-26

; PRIOR APPLICATION NUMBER: RU 2001117632

; PRIOR FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: RU 2001117633

; PRIOR FILING DATE: 2001-06-28

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 111

; TYPE: PRT

; ORGANISM: Escherichia coli

US-10-073-293A-6

Query Match 100.0%; Score 552; DB 5; Length 111;

Best Local Similarity 100.0%; Pred. No. 3,1e-59; Indels 0; Gaps 0;

Matches 111; Conservative 0; Mismatches 0;

QY 1 MSYEVLLGLLVGVNYCFRYPRLRVGNARPTKRGAVGILLDTIGIASICALLVVSTA 60

Db 1 MSYEVLLGLLVGVNYCFRYPRLRVGNARPTKRGAVGILLDTIGIASICALLVVSTA 60

QY 61 PEVMDTRFPVTLVGFVILGASFYKRSIIIPITLLSALAYGLAWKVMII 111

Db 61 PEVMDTRFPVTLVGFVILGASFYKRSIIIPITLLSALAYGLAWKVMII 111

RESULT 2

US-10-424-599-223037

; Sequence 223037, Application US/10424599

; Publication No. US20040031072A1

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 223037
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(400)
; OTHER INFORMATION: Clone ID: PAT_MRT3847_43431C.1.pep
US-10-424-599-223037

Query Match 15.2%; Score 84; DB 4; Length 223;
Best Local Similarity 29.8%; Pred. No. 0.11;
Matches 25; Conservative 16; Mismatches 21; Indels 22; Gaps 4;

Qy 39 VGILLDTIGIASICALLVSTAPEVMDTRRF-----VPTLVGFAVLGA--SFYKTRSIII 92
Db 132 VWITDKI-----LMLFSQDPEISHAREYCYILIPALFGHVLQALTRYFQTSWIP 184

Qy 93 PTLISA-----LAYGLAWK 107
Db 185 PMVFSITALCLHPICWGLVFKL 208

RESULT 3
US-10-732-923-23642
; Sequence 23642, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 23642
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(400)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-23642

Query Match 14.6%; Score 80.5; DB 5; Length 400;
Best Local Similarity 31.1%; Pred. No. 0.62;
Matches 32; Conservative 18; Mismatches 32; Indels 21; Gaps 8;

Qy 2 SYEVLLLG-LLVGV-ANYCFRYPRLRLRVGNARPTK-RGAVGILLDTIGIASICALLVVS 58
Db 19 SVMELILGRLLIGFCGLCTGFVP--MYIGVSPALRGAGF-TLNQLGI--VIGILVAQ 73

Qy 59 T-APEVMDTRRFVPTLVGFAVLGASFYKTRSIITLLSALA 100
Db 74 IFGLEIILGSEVLWPLVLGFT-----IIPALQSAA 104

RESULT 4
US-10-424-599-161652
; Sequence 161652, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
```

```
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 161652
; LENGTH: 1238
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1238)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_116990C.1.pep
US-10-424-599-161652

Query Match 14.4%; Score 79.5; DB 4; Length 1238;
Best Local Similarity 31.8%; Pred. No. 3.4;
Matches 21; Conservative 13; Mismatches 19; Indels 13; Gaps 3;

Qy 41 ILLDTIGIASICALLVSTAPEVMDTRRF-----VPTLVGFAVLGAS--FYKTRSIIPT 94
Db 96 IFVDKI-----LLLFQDPEISHAHEYCIYIPALYGFVAVLQCOIRYFQTOSMIFPM 148

Qy 95 LLSALA 100
Db 149 VFSSIA 154

RESULT 5
US-10-732-923-23556
; Sequence 23556, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 23556
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Canis familiaris
; OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-23556

Query Match 14.2%; Score 78.5; DB 5; Length 495;
Best Local Similarity 31.1%; Pred. No. 1.4;
Matches 32; Conservative 18; Mismatches 32; Indels 21; Gaps 8;

Qy 2 SYEVLLLG-LLVGV-ANYCFRYPRLRLRVGNARPTK-RGAVGILLDTIGIASICALLVVS 58
Db 116 SVMELILGRLLIGFCGLCTGFVP--MYIGVSPALRGAGF-TLNQLGI--VIGILVAQ 170

Qy 59 T-APEVMDTRRFVPTLVGFAVLGASFYKTRSIITLLSALA 100
Db 171 IFGLKVMINGTEELWPLLLGFT-----IIPAVLOSAA 201

RESULT 6
US-10-282-122A-63783
; Sequence 63783, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
```

```

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essee
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed -
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 62944
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-10-282-122A-62944

Query Match      14.0%; Score 77.5;
Best local similarity 24.8%; Pred. No. 1;
Matches 32; Conservative 20; Mismatchch
                                |||::
QY    9 GLLGVGVNVCFRY-----LPLRLH
DB   255 GALIGLSACTAFITANIDLTLDADLPMLQL
OV   56 VVSTAPRVWH-----DTPRPVDPTAVGR

```

[illegible]

/ APPLICANT: Cao, Jiongwei
 / APPLICANT: Hinkle, Gregory J.
 / APPLICANT: Slater, Steven S.
 / APPLICANT: Goldman, Barry C.
 / APPLICANT: Chen, Xianfeng
 / TITLE OF INVENTION: EXPRESSION OF MICROBIAL
 / TITLE OF INVENTION: PLANTS WITH IMPROVED F
 / FILE REFERENCE: 38-10(52052)B
 / CURRENT APPLICATION NUMBER: US/10/369,493
 / CURRENT FILING DATE: 2003-02-28
 / PRIOR APPLICATION NUMBER: US 60/360,039
 / PRIOR FILING DATE: 2002-02-21
 / NUMBER OF SEQ ID NOS: 47374
 / SEQ ID NO 3529
 / LENGTH: 308
 / TYPE: prt
 / ORGANISM: Neurospora crassa

PROTEINS IN PLANTS FOR PRODUCTION OF COPOLYMERS

US-10-369-493-3529

Query Match 13.9%; Score 77; DB 4; Length 308;
Best Local Similarity 39.8%; Pred. No. 1.2;
Matches 39; Conservative 5; Mismatches 42; Indels 12; Gaps 6;
Qy 6 LLLGLLVGVANVCFRYLPLRL-RVGNARPTKRGAVGILLDTI-GIAGICALLVVSTAPEV 63
Db 90 LVSSLAGVTSVFYTY-PLELIRVRLAFETKRGSSLSRSTIIRQIYSENALTVPKNAPAS 148
Qy 64 MHDTRRFVPTLVGPAVLGASPYKTRSIITPLLSALAY 101
Db 149 AH-----APALIPRTGL-ANFYRGFS---PTLLQMLPY 177

RESULT 9

US-10-732-923-23831
; Sequence 23831, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 23831
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Ovis aries
US-10-732-923-23831

Query Match 13.9%; Score 76.5; DB 5; Length 161;
Best Local Similarity 34.1%; Pred. No. 0.61;
Matches 29; Conservative 17; Mismatches 30; Indels 9; Gaps 7;
Qy 2 SYEVLILG-LLVGV-ANVCFRYLPLRLRVGNARPTK-RGAVGILLDTIGIASICALLVWS 58
Db 29 SVEMLILGRLLIIGLFCGLCTGFVP--MYGISEPTALRGAFG-TLNQLGI--VIGILVAQ 83
Qy 59 T-APEVMHDTRRFVPTLVGPAVLGA 82
Db 84 IFGLKVILGTEDLWPLLGLFTILPA 108

RESULT 10

US-09-981-947A-5
; Sequence 5, Application US/09981947A
; Patent No. US20020164578A1
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; Weng, Kun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/981,947A
; FILING DATE: 18-Oct-2001
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/031,392
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-981-947A-5

Query Match 13.9%; Score 76.5; DB 3; Length 494;
Best Local Similarity 34.1%; Pred. No. 2.5;
Matches 29; Conservative 17; Mismatches 30; Indels 9; Gaps 7;
Qy 2 SYEVLILG-LLVGV-ANVCFRYLPLRLRVGNARPTK-RGAVGILLDTIGIASICALLVWS 58
Db 116 SVEMLILGRLLIIGLFCGLCTGFVP--MYGISEPTALRGAFG-TLNQLGI--VIGILVAQ 170
Qy 59 T-APEVMHDTRRFVPTLVGPAVLGA 82
Db 171 IFGLKVILGTEDLWPLLGLFTILPA 195

RESULT 11

US-10-732-923-23560
; Sequence 23560, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 23560
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-732-923-23560

Query Match 13.9%; Score 76.5; DB 5; Length 494;
Best Local Similarity 34.1%; Pred. No. 2.5;
Matches 29; Conservative 17; Mismatches 30; Indels 9; Gaps 7;
Qy 2 SYEVLILG-LLVGV-ANVCFRYLPLRLRVGNARPTK-RGAVGILLDTIGIASICALLVWS 58
Db 116 SVEMLILGRLLIIGLFCGLCTGFVP--MYGISEPTALRGAFG-TLNQLGI--VIGILVAQ 170
Qy 59 T-APEVMHDTRRFVPTLVGPAVLGA 82
Db 171 IFGLKVILGTEDLWPLLGLFTILPA 195

RESULT 12

US-10-732-923-23639
; Sequence 23639, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923


```
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 23639
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Ovis aries
US-10-732-923-23639

Query Match      13.9%; Score 76.5; DB 5; Length 494;
Best Local Similarity 34.1%; Pred. No. 2.5;
Matches 29; Conservative 17; Mismatches 30; Indels 9; Gaps 7;

QY  2 SYEVLLIG-LLGV-ANYCFRYPRLRLRVGNARPTK-RGAVGILLDTIGIASICALVWS 58
Db  116 SVEMLILGRVLIGLFCGLCTGFVP--MYIGISPTALRGAFG-TLNQLGI--VIGILVAQ 170

QY  59 T-APEVMHDTRRFVPTLVGFVAVLGA 82
Db  171 IFGLKVLGTEDLWPLLLIGFTILPA 195

RESULT 13
US-10-369-493-12445
; Sequence 12445, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12445
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(500)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-12445

Query Match      13.9%; Score 76.5; DB 4; Length 500;
Best Local Similarity 25.2%; Pred. No. 2.5;
Matches 33; Conservative 20; Mismatches 45; Indels 33; Gaps 5;

QY  9 GLLVGVANYCFRYPRLRLRVGNARPTKRGAVGILLDTIGIASICALVWSTAPFVMDTR 68
Db  20 GLFVGIGAYLRDAGPLSLGLY----LIWGIAFILPILNSVGEMCAYLPIRGS--IFELAA 74

QY  69 RFVPTLVGFVAVLGSFYK---TRS-----IIIPFLSALAYGL----- 103
Db  75 RYVDPARFGFAMXGSPARFVQGRSDNIRLGDWVVFYAGLMVLCVCTEYSAVAFIMDYQIDV 134

QY  104 ---AWKVMAIL 111
Db  135 NPAAWVAMW 145

RESULT 14
US-10-092-900A-314
; Sequence 314, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsbrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 314
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-314

Query Match      13.8%; Score 76; DB 4; Length 509;
Best Local Similarity 32.1%; Pred. No. 3;
Matches 27; Conservative 17; Mismatches 34; Indels 6; Gaps 5;

QY  2 SYEVLLIG-LLGV-ANYCFRYPRLRLRVGNARPTK-RGAVGILLDTIGIASICALVWS 58
Db  119 SVEMLILGRVLIGLFCGLCTGFVP--MYIGISPTALRGAFG-TLNQLGIVIGILVAQVI 175

QY  59 TAPEVMHDTRRFVPTLVGFVAVLGA 82
```

Db 176 FGLEILGSEELWVPVLGFTILPA 199

RESULT 15

US-10-437-963-107875
 ; Sequence 107875, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 107875
 ; LENGTH: 538
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_12184C.1.pap
 US-10-437-963-107875

Query Match 13.8%; Score 76; DB 4; Length 538;
 Best Local Similarity 26.8%; Pred No. 3.2;
 Matches 30; Conservative 15; Mismatches 27; Indels 40; Gaps 6;
 QY 6 LLLGLLVGVNYCFRFLRLRVGNARPTKRGAVGILLD-----TIGIASICAL- 54
 Db 84 LWVGLVGVPSY---YL-----AGSLVDLGMSALQGVATVAFANLVLV 124
 QY 55 -LVVSTAPEVMDTRFVPTLVGFVILGASFYKTRSIIPILLSMAYGLAW 105
 Db 125 SLVLTAPAVTHG-----LPFPVLARAAFGVGRGAHLPAVIRALV-GCGW 167

Search completed: February 15, 2006, 12:29:34
 Job time : 72.4017 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 12:25:52 ; Search time 6.54775 Seconds
(without alignments)
240.922 Million cell updates/sec

Title: US-10-073-293A-6
Perfect score: 552
Sequence: 1 MSTEVLLGLLVGVNYCFR.....IPTLLMALVGLAWKNVNI 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 107799 seqs, 14211699 residues

Total number of hits satisfying chosen parameters: 107799

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pdb.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb.*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb.*
5: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb.*
6: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pdb.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pdb.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	13.4	202	7	US-11-082-389-364
2	72.5	13.1	499	6	US-10-821-234-1106
3	67	12.1	1034	6	US-10-392-234A-30
4	66.5	12.0	177	6	US-10-793-626-2532
5	66.5	12.0	203	6	US-10-453-372-890
6	64.5	11.7	185	6	US-10-453-372-896
7	64.5	11.7	194	6	US-10-453-372-894
8	64.5	11.7	218	6	US-10-453-372-892
9	64	11.6	272	7	US-11-098-686-10661
10	64	11.6	332	6	US-10-454-437-28
11	64	11.6	1037	6	US-10-392-234A-36
12	63.5	11.5	332	6	US-10-511-538-79
13	63.5	11.5	551	6	US-10-055-877-276
14	63.5	11.5	551	6	US-10-055-877-277
15	63	11.4	1048	6	US-10-392-234A-14
16	63	11.4	1048	6	US-10-392-234A-20
17	62.5	11.3	485	6	US-10-467-657-326
18	62	11.2	314	6	US-10-511-538-70
19	62	11.2	1049	6	US-10-392-234A-12
20	61.5	11.1	434	6	US-10-821-234-1680
21	61.5	11.1	520	7	US-11-098-686-11110
22	61	11.1	325	6	US-10-510-101-156
23	61	11.1	983	6	US-10-511-538-91
24	61	11.1	1048	6	US-10-392-234A-18
25	60.5	11.0	294	6	US-10-793-626-1688

26	60.5	11.0	1046	6	US-10-392-234A-16
27	59.5	10.8	228	7	US-11-210-316-16
28	59.5	10.8	295	6	US-10-858-730-112
29	59.5	10.8	480	7	US-11-080-991-76
30	59.5	10.8	518	6	US-10-821-234-1165
31	59	10.7	244	6	US-10-793-626-134
32	59	10.7	352	6	US-10-793-626-944
33	59	10.7	374	7	US-11-228-364-4
34	59	10.7	654	7	US-11-120-308-52
35	59	10.7	895	7	US-11-098-686-10311
36	58.5	10.6	233	6	US-10-858-730-234
37	58.5	10.6	233	7	US-11-055-822-54
38	58.5	10.6	454	6	US-10-055-877-204
39	58.5	10.6	747	7	US-11-210-316-2
40	58.5	10.6	1061	7	US-11-059-814-18
41	58	10.5	311	6	US-10-980-388-111
42	58	10.5	311	6	US-10-980-388-113
43	58	10.5	530	6	US-10-980-388-62
44	57.5	10.4	303	7	US-11-052-554A-321
45	57.5	10.4	488	6	US-10-055-877-231

ALIGNMENTS

RESULT 1
US-11-082-389-364
; Sequence 364, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE TRANSPORT
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131CPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 364
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-364

Query Match 13.4%; Score 74; DB 7; Length 202;
Best Local Similarity 27.8%; Pred. NO. 0.43;
Matches 30; Conservative 18; Mismatches 38; Indels 22; Gaps 5;

Qy 5 VLLGLLVGVANVCYRYLPLRLR-----VGNARPT--KRGAVGILLDTIGIASICAL 54
 Db 60 IMLGLVGLAWDYQIFLVTRREGFTKGTAGNATSGFKHGA-----RVVTRAAAL 111
 Qy 55 LVVSTAPEVMDHTRRFVPTLVGFVGLGASFYK---TRSIITPTLLSAL 99
 Db 112 IMVSVFAAFIAQDMAFIKTM-GFALAVAVFFDAFVVRMMIIPATMFL 158

RESULT 2

US-10-821-234-1106
 ; Sequence 1106, Application US/10821234
 ; Publication No. US20050255114A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Labat, Ivan
 ; APPLICANT: Stache-Crain, Birgit
 ; APPLICANT: Andarmani, Susan
 ; APPLICANT: Tang, Y. Tom
 ; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
 ; FILE REFERENCE: 821A
 ; CURRENT APPLICATION NUMBER: US/10/821,234
 ; CURRENT FILING DATE: 2004-04-07
 ; PRIOR APPLICATION NUMBER: US 60/462,047
 ; PRIOR FILING DATE: 2003-04-07
 ; NUMBER OF SEQ ID NOS: 1704
 ; SOFTWARE: dt_seq_genes Version 1.0
 ; SEQ ID NO 1106
 ; LENGTH: 499
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-821-234-1106

Query Match 13.1%; Score 72.5; DB 6; Length 499;
 Best Local Similarity 32.9%; Pred. No. 1.7;
 Matches 28; Conservative 17; Mismatches 31; Indels 9; Gaps 7;

Qy 2 SYEVLLLG-LLVGV-ANYCFRYLPLRLRVGNARPTK-RGAVGILLDTIGIASICALLVVS 58
 Db 119 SVEMLILGRLVIGFCGLCTGFVP--WYIGISPTALRGAGF-TLNQGI--VVGILVAQ 173
 Qy 59 T-APEVMDHTRRFVPTLVGFVGLGA 82
 Db 174 IFGLEFILGSEELWPLILGFTILPA 198

RESULT 3

US-10-392-234A-30
 ; Sequence 30, Application US/10392234A
 ; Publication No. US2005025538A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pharmacia and Upjohn Corporation
 ; APPLICANT: Buxser, Steven
 ; APPLICANT: Poole, Keith
 ; APPLICANT: Decker, Douglas
 ; APPLICANT: Xiazhi Li
 ; TITLE OF INVENTION: Method for Screening for acrAB Transporter Family Inhibitors
 ; FILE REFERENCE: 6206
 ; CURRENT APPLICATION NUMBER: US/10/392,234A
 ; CURRENT FILING DATE: 2003-03-17
 ; PRIOR APPLICATION NUMBER: US 60/364,935
 ; PRIOR FILING DATE: 2002-03-15
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 30
 ; LENGTH: 1034
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-392-234A-30

Query Match 12.1%; Score 67; DB 6; Length 1034;
 Best Local Similarity 30.4%; Pred. No. 15;
 Matches 21; Conservative 15; Mismatches 25; Indels 8; Gaps 1;

Qy 42 LLDITIGIASICALVWSTAPEVMDHTRRFVPTLVGFVGLGASFYKTRSIITPTLLSALAY 101
 Db 930 LLTTIGLSAKNAIILVFEKDLMEKE-----GKGVVEATLMAVRMLRPLMTSLAP 981
 Qy 102 GLAWKVMAI 110
 Db 982 ILGVLPLAI 990

RESULT 4

US-10-793-626-2532
 ; Sequence 2532, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMMERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PU3480US
 ; CURRENT APPLICATION NUMBER: US/10/793,626
 ; CURRENT FILING DATE: 2004-03-04
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2532
 ; LENGTH: 177
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: amino acid sequence
 US-10-793-626-2532

Query Match 12.0%; Score 66.5; DB 6; Length 177;
 Best Local Similarity 28.3%; Pred. No. 2.3;
 Matches 32; Conservative 16; Mismatches 48; Indels 17; Gaps 5;
 Qy 6 LLLGLLVGVANVCYRYLPLRLRVGNARPT-----KRGAVGILLDTIGIASICALLVVST 59
 Db 30 VFLGIFIGYAGY---YL-LRKNFLAMPSLIEQGFSGELGIALSAVSIAYGFSKVMGT 85
 Qy 60 APE-----VMDHTRRFVPTLVGFVGLGASFYKTRSIITPTLLSALA----YGLAW 105
 Db 86 VSDRSNARMFLTLGLVLTATINLLGLFPFTSSITIMFIMFLVGVFGQMGW 138

RESULT 5

US-10-453-372-890
 ; Sequence 890, Application US/10453372
 ; Publication No. US2006000323A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alcobrook, et al.
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
 ; FILE REFERENCE: 21402-589 A
 ; CURRENT APPLICATION NUMBER: US/10/453,372
 ; CURRENT FILING DATE: 2003-06-03
 ; PRIOR APPLICATION NUMBER: 09/789390
 ; PRIOR FILING DATE: 2001-02-23
 ; PRIOR APPLICATION NUMBER: 60/185967
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: 09/823187
 ; PRIOR FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: 60/195792
 ; PRIOR FILING DATE: 2000-03-10
 ; PRIOR APPLICATION NUMBER: 09/839446
 ; PRIOR FILING DATE: 2001-03-19
 ; PRIOR APPLICATION NUMBER: 60/199476
 ; PRIOR FILING DATE: 2000-03-25
 ; PRIOR APPLICATION NUMBER: 09/863776
 ; PRIOR FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: 60/208263
 ; PRIOR FILING DATE: 2000-05-31
 ; PRIOR APPLICATION NUMBER: 09/939398
 ; PRIOR FILING DATE: 2001-08-24

```

; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 890
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-890

Query Match      12.0%; Score 66.5; DB 6; Length 203;
Best Local Similarity 27.5%; Pred. No. 2.7;
Matches 25; Conservative 16; Mismatches 39; Indels 11; Gaps 4;

Qy  5  VLLGLGLGVGANYCFRYLPRL-LRVGNARPTKRGAVGILLDTIGI-ASICALIIVSTAPE 62
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  83  LMVAVLLGFVAMVLSVVGKCTRVGDSNPIAKGRVAIAGALFILAGLCTLTAVSWYAT 142
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy  63  VMHD---TRRPVPTL-----VGFAVLGASF 84
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  143  LVTQEPFNPEFGPALFVGWASAGLAVLGGSF 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-10-453-372-896
; Sequence 896, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
; APPLICANT: Alsebrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453.372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 896
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-896

Query Match      11.7%; Score 64.5; DB 6; Length 185;
Best Local Similarity 27.3%; Pred. No. 4;
Matches 27; Conservative 15; Mismatches 38; Indels 19; Gaps 5;

Qy  5  VLLGLGLGVGANYCFRYLPRL-LRVGNARPTKRGAVGILLDTIGI-ASICALIIVS---- 58
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  83  LMVAVLLGFVAMVLSVVGKCTRVGDSNPIAKGRVAIAGALFILAGLCTLTAVSWYAT 142
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy  59  -----TAPEVMHDTR-RFVPTL-----VGFAVLGASF 84
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 143 LVTQEFNFTPNVARYEFGPALFVGWASAGLAVLGGSF 181

RESULT 7.

US-10-453-372-894

; Sequence 894, Application US/10453372

; Publication No. US2006000323A1

; GENERAL INFORMATION:

; APPLICANT: Alsbrook, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS

; FILE REFERENCE: 21402-589 A

; CURRENT APPLICATION NUMBER: US/10/453,372

; CURRENT FILING DATE: 2003-06-03

; PRIOR APPLICATION NUMBER: 09/789390

; PRIOR FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: 60/185967

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 09/823187

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/195792

; PRIOR FILING DATE: 2000-03-10

; PRIOR APPLICATION NUMBER: 09/839446

; PRIOR FILING DATE: 2001-03-19

; PRIOR APPLICATION NUMBER: 60/199476

; PRIOR FILING DATE: 2000-03-25

; PRIOR APPLICATION NUMBER: 09/863776

; PRIOR FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: 60/208263

; PRIOR FILING DATE: 2000-05-31

; PRIOR APPLICATION NUMBER: 09/939398

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: 60/227800

; PRIOR FILING DATE: 2000-08-25

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 1609

; SOFTWARE: Curaseqlist version 0.1

; SEQ ID NO 894

; LENGTH: 194

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-453-372-894

Query Match 11.7%; Score 64.5; DB 6; Length 194;

Best Local Similarity 27.3%; Pred.No. 4.2;

Matches 27; Conservative 15; Mismatches 38; Indels 19; Gaps 5;

QY 5 VLLGLLVGVANVCFRYLPLR-LRVGNAPRTKGGAVILLDTIGI-ASICALLVWS----- 58

Db 63 LHMVAVLLGFGVAMLVSVWGKCTRVGDSNP IAKGRVIAIAGGALFILAGLCTLTAVSWYAT 122

QY 59 -----TAPVMDHTR-RFVPTL-----VGFVAVLGASF 84

Db 123 LVTQEFNFTPNVARYEFGPALFVGWASAGLAVLGGSF 161

RESULT 8

US-10-453-372-892

; Sequence 892, Application US/10453372

; Publication No. US2006000323A1

; GENERAL INFORMATION:

; APPLICANT: Alsbrook, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS

; FILE REFERENCE: 21402-589 A

; CURRENT APPLICATION NUMBER: US/10/453,372

; CURRENT FILING DATE: 2003-06-03

; PRIOR APPLICATION NUMBER: 09/789390

; PRIOR FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: 60/185967

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 09/823187

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/195792

; PRIOR FILING DATE: 2000-03-10

Db 143 LVTQEFNFTPNVARYEFGPALFVGWASAGLAVLGGSF 181

RESULT 7.

US-10-453-372-894

; Sequence 894, Application US/10453372

; Publication No. US2006000323A1

; GENERAL INFORMATION:

; APPLICANT: Alsbrook, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS

; FILE REFERENCE: 21402-589 A

; CURRENT APPLICATION NUMBER: US/10/453,372

; CURRENT FILING DATE: 2003-06-03

; PRIOR APPLICATION NUMBER: 09/789390

; PRIOR FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: 60/185967

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 09/823187

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/195792

; PRIOR FILING DATE: 2000-03-10

; PRIOR APPLICATION NUMBER: 09/839446

; PRIOR FILING DATE: 2001-03-19

; PRIOR APPLICATION NUMBER: 60/199476

; PRIOR FILING DATE: 2000-03-25

; PRIOR APPLICATION NUMBER: 09/863776

; PRIOR FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: 60/208263

; PRIOR FILING DATE: 2000-05-31

; PRIOR APPLICATION NUMBER: 09/939398

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: 60/227800

; PRIOR FILING DATE: 2000-08-25

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 1609

; SOFTWARE: Curaseqlist version 0.1

; SEQ ID NO 894

; LENGTH: 194

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-453-372-894

Query Match 11.7%; Score 64.5; DB 6; Length 194;

Best Local Similarity 27.3%; Pred.No. 4.2;

Matches 27; Conservative 15; Mismatches 38; Indels 19; Gaps 5;

QY 5 VLLGLLVGVNYCFRFLPLR-LRVGNARPTKRGAVILLDTIGI-ASICALLVWS----- 58

Db 63 LMVAVLLGFGVAVMLSVWGKCTRVGDSNP IAKGRVIAIAGGALFILAGLCTLTAVSWYAT 122

QY 59 -----TAPVMDHTR-RFVPTL-----VGFVAVLGASF 84

Db 123 LVTQEFNFTPNVARYEFGPALFVGWASAGLAVLGGSF 161

RESULT 8

US-10-453-372-892

; Sequence 892, Application US/10453372

; Publication No. US2006000323A1

; GENERAL INFORMATION:

; APPLICANT: Alsbrook, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS

; FILE REFERENCE: 21402-589 A

; CURRENT APPLICATION NUMBER: US/10/453,372

; CURRENT FILING DATE: 2003-06-03

; PRIOR APPLICATION NUMBER: 09/789390

; PRIOR FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: 60/185967

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 09/823187

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/195792

; PRIOR FILING DATE: 2000-03-10

```
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Curaseq1ist version 0.1
; SEQ ID NO 892
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-892

Query Match      11.7%; Score 64.5; DB 6; Length 218;
Best Local Similarity 27.3%; Pred. No. 4.8;
Matches 27; Conservative 15; Mismatches 38; Indels 19; Gaps 5;

Qy  5 VLLIGLLGVNVCYFVPLR-LVGNARPTKRGAVGILLDTIGI-ASICALLVVS----- 58
Db  87 LNVVAVLLGFMVVLVSVVGVGKTRVGSNPTAKGRVAIAGALFILAGLCTLTAVSWYAT 146
Qy  59 -----TAPEVMHDTIR-RFVPTL-----VGFAVLGASF 84
Db  147 LVTOEFNPTFVNARVEFGFALFVGWASAGLAVLGGSF 185

RESULT 9
US-11-098-686-10661
; Sequence 10661, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10661
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10661

Query Match      11.6%; Score 64; DB 7; Length 272;
Best Local Similarity 24.3%; Pred. No. 7;
Matches 28; Conservative 23; Mismatches 48; Indels 16; Gaps 6;

Qy  1 MSYEVLLGLLVGNVNYCFRY---LPLRLRVGNARPTKRGAVGILLDTIGIATICALIVV 57
Db  1 MSIELLYTGVLQGLTLAIATFGIMIPFRF-LNFPDLTAEGAY-----PLG-GAVCASLMV 53
Qy  58 STAPEVMHDTIRRFVPTLVGFAVLGAS-----FVKTRSIITPTLLGALAYGLAWKVM 108
Db  54 ANMFQILAWMAGIMAG--GLLAVGTQSVALRLKVNLLAGILLSTWAYSINLRIM 106

RESULT 10
US-10-454-437-28
; Sequence 28, Application US/10454437
```

```
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 199332125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 28
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-28

Query Match      11.6%; Score 64; DB 6; Length 332;
Best Local Similarity 26.1%; Pred. No. 8.7;
Matches 30; Conservative 13; Mismatches 38; Indels 34; Gaps 7;

Qy  31 ARPTKRGAVGI-----LLD-TIGIASIC-----ALLVVSTAPEV---MHDTRRFVPTLV- 75
Db  58 AAPLAGADAIFNORTIIDVAIAFVFCFASAIYLVNDARDVDREHPTRKRPPIAAG 117
Qy  76 -----GFAV-----LGASFYKTRSIITPTLLSA-----LAYGLAWKVM 111
Db  118 VLPVGMAYGMAVALIALSIGLSFLATDGVALACVIGVYIALQLGVCFGWKHPVI 172

RESULT 11
US-10-392-234A-36
; Sequence 36, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxser, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiaozhi Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 1037
```

```

; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Cahterine
; APPLICANT: Eisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shinkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 276
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-055-877-276

Query Match 11.5%; Score 63.5; DB 6; Length 551;
Best Local Similarity 25.6%; Pred. No. 17;
Matches 30; Conservative 15; Mismatches 37; Indels 35; Gaps

Qy 4 EYLLGLGL-LGVANYCFRYLP-----LRLRVGNARPTKRGAVGILLDTIGIASICALL
Db 163 KVLILNYLQTAVSGTCAAYAPNTYVCVFLLSG-----MSLASIAINCMTL;

Qy 56 VVSTAPEVMHDTFRPFTLVGFAVLGASFYKTRSIIFTLLSALAYGLA-KVVMAL 111
Db 210 NVEMWP--IH-TRAYVGLTIGVYVSLGQF-----LLAGIAVAVPHWRHLQLV 253

RESULT 14
US-10-055-877-277
; Sequence 277, Application US/10055877
; Publication No. US20050288241A1

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:05:59 ; Search time 2814.6 Seconds
(without alignments)
2241.749 Million cell updates/sec

Title: US-10-073-293a-6

Perfect score: 552

Sequence: 1 MSFEVLLGLLVGVNYCFR.....IPTLLSALAYGLAWKVMAII 111

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/absa/ABSSWEB spool/US10073293/runat_15022006_120734_963/app_query_fasta.1
-DB=GenEmbl -QPM=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs04
-USER=US10073293 @CGN 1 1 5142 @runat_15022006_120734_963 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_ov.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_ets.*
- 11: gb_ey.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	552	100.0	336	6	BD177949 Process f
2	552	100.0	336	6	CS078099 Sequence
3	552	100.0	336	6	CS078115 Sequence

4	552	100.0	336	6	AX534670	AX534670 Sequence
5	552	100.0	19150	1	D90891	D90891 E.coli geno
6	552	100.0	110000	1	U00096_27	Continuation (28 o
7	552	100.0	110000	1	U00096_28	Continuation (29 o
8	548	99.3	110000	1	AE005674_27	Continuation (28 o
9	548	99.3	110000	1	BA000007_35	Continuation (36 o
10	548	99.3	290380	1	AE016987	AE016987 Shigella
11	548	99.3	305325	1	AE016765	AE016765 Escherich
12	544	98.6	110000	1	AE005174_35	Continuation (36 o
13	544	98.6	110000	1	AE005174_36	Continuation (37 o
14	403	73.0	372	6	AR386318	AR386318 Sequence
15	313.5	56.8	3619	1	EC410307	Au410307 Erwinia c
16	309.5	56.1	110000	1	EX936399_10	Continuation (11 o
17	309.5	56.1	110000	1	EX950851_39	Continuation (40 o
18	308.5	55.9	10733	1	AE013695	AE013695 Yersinia
19	308.5	55.9	220050	1	AJ414156	AJ414156 Yersinia
20	308.5	55.9	290002	1	AE017129	AE017129 Yersinia
21	295	53.4	328	1	ECU19993	EU9993 Escherichia
22	271.5	49.2	243072	6	AX770904	AX770904 Sequence
23	271.5	49.2	349318	1	EX571863	EX571863 Phototab
24	182.5	33.1	369	6	AR320188	AR320188 Sequence
25	164	29.7	100	6	CQ000312	CQ000312 Sequence
26	160	29.0	100	6	CQ000311	CQ000311 Sequence
27	155	28.1	100	6	CQ000313	CQ000313 Sequence
28	147	26.6	114080	14	AP007521	AP007521 Lotus cor
29	87	15.8	307337	1	EX842656	EX842656 Bdellovib
30	86.5	15.7	90018	14	AC017853	AC017853 Drosophil
31	86.5	15.7	167026	2	AC010011	AC010011 Drosophil
32	86.5	15.7	195757	2	AC010009	AC010009 Drosophil
33	86.5	15.7	249021	2	AE003596	AE003596 Drosophil
34	86	15.6	1851	5	BC097869	BC097869 Xenopus 1
35	86	15.6	268356	1	TACID5	AL445067 Thermopla
36	85	15.4	233445	14	AC129682	AC129682 Rattus no
37	84.5	15.3	921	15	AY857266	AY857266 Hypocrea
38	83	15.0	632	15	AF348090	AF348090 Trichoder
39	82.5	14.9	19557	13	AY029400	AY029400 Campolet
40	82	14.9	951	4	AY355763	AY355763 Canis fam
41	82	14.9	110000	1	BA000012_62	Continuation (63 o
42	82	14.9	110000	14	AP006495	AP006495 Cyanidios
43	82	14.9	303250	1	BX294143	BX294143 Pirellula
44	81.5	14.8	483	4	OAUS9030	U9030 Ovis aries
45	81.5	14.8	1793	4	AY033938	AY033938 Bos tauri

ALIGNMENTS

RESULT 1

BD177949

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD177949 Process for producing L-amino acid using escherichia. linear PAT 16-APR-2003

BD177949

BD177949, 1 GI:30015212

JP 2002300874-A/4.

Escherichia coli

Escherichia coli

1 (bases 1 to 336)

Tabolina, E.A., Rybak, K.V., Khourges, E.M., Voroshilova, E.B. and

Gusyatiner, M.M.

Patent: JP 2002300874-A 4 15-OCT-2002;

OS Escherichia coli

PN JP 2002300874-A/4

PD 15-OCT-2002

PR 13-FEB-2001 RU 2001103865, 26-FEB-2001 RU 2001104998 PR

26-FEB-2001 RU 2001104999, 28-JUN-2001 RU 2001117632 PR

28-JUN-2001 RU 2001117633

PI EKATERINA ALEKSANDROVNA TABOLINA, KONSTANTIN VYACHESLAVOVICH

PI RYBAK,

PI EVGENI MOISEVICH KHOURGES, ELVIRA BORISOVNA VOROSHILOVA, PI

MIKHAIL MAROVICH GUSYATINER
PC C12N1/21, C12P13/04, C12P13/06, C12P13/08, C12P13/12, PC
C12P13/24//
PC (C12N1/21, C12R1:19), (C12P13/04, C12R1:19), (C12P13/06, C12R1:19),
(C12P13/08, C12R1:19), (C12P13/12, C12R1:19), (C12P13/12, C12R1:19), PC
(C12P13/24, C12R1:19)
CC Process for producing L-amino acid using *Escherichia FH* Key
Location/Qualifiers
FT CDS Location/Qualifiers
(1)..(333).

FEATURES

source
1..336
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"

ORIGIN

Alignment Scores:
Pred. No.: 1.16e-50 Length: 336
Score: 552.00 Matches: 111
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-073-293A-6 (1-111) x BD177949 (1-336)

Qy 1 MetSerTyrGluValLeuLeuGlyLeuValGlyValAlaAAsnTyrCysPheArg 20
Db 1 ATGAGCTATGAGGTTCTGCTGGGTACTAGTTGGCGTGGCGAATTATTGCTTCGCG 60
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 61 TATTGGCGCTGGCGCTCGGTGGGTAAATGCCCGCCCAACAAACGTCGCGGTAGGT 120
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAla 60
Db 121 ATTTTGTCCGACCATTTGGCATGCCCTCGATATGGCTCTGCTGGTGTCTCTACCGCA 180
Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 181 CCAGAGTGATGACGATACACGCGGTTTCGTGGCCACGCTGCGGTCTCGCGGTACTG 240
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 241 GTGCGCAGTTCTTATAAACACGAGCATTTATCATCCCAACACTGCTTAGTGGCGTGGCC 300

RESULT 2
CS078099
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..336
/organism="Escherichia coli"
/mol_type="unassigned DNA"

1 Tabolina,E.A., Rybak,K.V., Kfourges,E.M., Voroshilova,E.B. and
Gusyatiner,M.M.
Method for producing L-amino acid using bacteria belonging to the
genus *Escherichia*
Patent: EP 1526179-A 5 27-APR-2005;
Ajinomoto Co., Inc. (JP)
Location/Qualifiers
1..336
/organism="Escherichia coli"
/mol_type="unassigned DNA"

CDS

/db_xref="taxon:562"
1..336
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAI94224.1"
/db_xref="GI:63093068"
/translation="MSYEVLLGLLVGVANVCFRFLPLRLRVGNARPTKRGAVGILLD
TIGIASICALVVSTAPEVMHDTFRFPTLVGFVAVLGASFYKTRSIITLLSALAYG
LAWKVMAIL"

ORIGIN

Alignment Scores:
Pred. No.: 1.16e-50 Length: 336
Score: 552.00 Matches: 111
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-073-293A-6 (1-111) x CS078099 (1-336)

Qy 1 MetSerTyrGluValLeuLeuGlyLeuValGlyValAlaAAsnTyrCysPheArg 20
Db 1 ATGAGCTATGAGGTTCTGCTGGGTACTAGTTGGCGTGGCGAATTATTGCTTCGCG 60
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 61 TATTGGCGCTGGCGCTCGGTGGGTAAATGCCCGCCCAACAAACGTCGCGGTAGGT 120
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAla 60
Db 121 ATTTTGTCCGACCATTTGGCATGCCCTCGATATGGCTCTGCTGGTGTCTCTACCGCA 180
Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 181 CCAGAGTGATGACGATACACGCGGTTTCGTGGCCACGCTGCGGTCTCGCGGTACTG 240
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 241 GTGCGCAGTTCTTATAAACACGAGCATTTATCATCCCAACACTGCTTAGTGGCGTGGCC 300
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 301 TATGGGCTCGCTGGAAAGTGATGGCGATTATA 333

RESULT 3

CS078115
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..336
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

1 Tabolina,E.A., Rybak,K.V., Kfourges,E.M., Voroshilova,E.B. and
Gusyatiner,M.M.
Method for producing L-amino acid using bacteria belonging to the
genus *Escherichia*
Patent: EP 1526181-A 5 27-APR-2005;
Ajinomoto Co., Inc. (JP)
Location/Qualifiers
1..336
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

CDS

CS078115
Sequence 5 from Patent EP1526181.
CS078115
CS078115.1 GI:63093083

Escherichia coli

Escherichia coli

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Escherichia.

1

Tabolina,E.A., Rybak,K.V., Kfourges,E.M., Voroshilova,E.B. and

Gusyatiner,M.M.

Method for producing L-amino acid using bacteria belonging to the

genus *Escherichia*

Patent: EP 1526181-A 5 27-APR-2005;

Ajinomoto Co., Inc. (JP)

Location/Qualifiers

1..336

/organism="Escherichia coli"

/mol_type="unassigned DNA"

/db_xref="taxon:562"

1..336

/note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CAI94228.1"
/db_xref="GI:63093084"
/translation="MSYEVLILGLLVGVANYCFRPLRLRVGNARPTKRGAVGILLD
TIGIASICALLVSTAPEVMDTRRFVTLVGVAVLGAFYKTRSIITLLSALAYG
LAWKVMII"

ORIGIN

Alignment Scores:

Pred. No.: 1.16e-50 Length: 336
Score: 552.00 Matches: 111
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-073-293A-6 (1-111) x CS078115 (1-336)

Qy 1 MetSerTyrGluValLeuLeuGlyValLeuValGlyValAlaAseTyrCysPheArg 20
Db 1 ATGAGCTATGAGGTTCTGCTGTGGGTTACTAGTTGGCGTGGCGAATTATTGCTCCGC 60
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 61 TATTGGCGCTGCCTCGCTGGTGGTAATGCCCGCCCAACCAACGTCGCGCGTAGGT 120
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAla 60
Db 121 ATTTTGTCTGACACCAATTGGCATCGCTCGATATGCGTCTGCTGGTGTCTCTACCGCA 180
Qy 61 ProGluValMetHisAspThrArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 181 CCAGAAAGTGATGCACGATACACGCCGTTTCGTGCCACCGCTGCTGCGGTACTG 240
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 241 GTTGCCAGTTCTATAAACAACGCGCATTTATCATCCACACATGCTTAGTGCGCTGCC 300
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 301 TATGGCTCGCTGGAAGTGATGGCGATTATA 333

RESULT 4

AX534670
LOCUS AX534670 336 bp DNA linear PAT 22-NOV-2002
DEFINITION Sequence 5 from Patent EP1239041.
ACCESSION AX534670
VERSION AX534670.1 GI:25261076

KEYWORDS

SOURCE

ORGANISM

Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE

AUTHORS

Tabolina,E.A., Rybak,K.V., Khourges,E.M., Voroshilova,E.B. and

Gusyatiner,M.M.

Method for producing l-amino acid using bacteria belonging to the

genus escherichia

Patent: EP 1239041-A 5 11-SEP-2002;

Ajinomoto Co., Inc. (JP)

FEATURES

source

1. .336

Location/Qualifiers

/organism="Escherichia coli"

/mol_type="unassigned DNA"

/db_xref="taxon:562"

1. .336

/note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CAD57226.1"

/db_xref="GI:25261077"

/translation="MSYEVLILGLLVGVANYCFRPLRLRVGNARPTKRGAVGILLD
TIGIASICALLVSTAPEVMDTRRFVTLVGVAVLGAFYKTRSIITLLSALAYG
LAWKVMII"

CDS

ORIGIN

Alignment Scores:

Pred. No.: 1.16e-50 Length: 336
Score: 552.00 Matches: 111
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-073-293A-6 (1-111) x AX534670 (1-336)

Qy 1 MetSerTyrGluValLeuLeuGlyValLeuValGlyValAlaAseTyrCysPheArg 20
Db 1 ATGAGCTATGAGGTTCTGCTGTGGGTTACTAGTTGGCGTGGCGAATTATTGCTCCGC 60
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 61 TATTGGCGCTGCCTCGCTGGTGGTAATGCCCGCCCAACCAACGTCGCGCGTAGGT 120
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAla 60
Db 121 ATTTTGTCTGACACCAATTGGCATCGCTCGATATGCGTCTGCTGGTGTCTCTACCGCA 180
Qy 61 ProGluValMetHisAspThrArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 181 CCAGAAAGTGATGCACGATACACGCCGTTTCGTGCCACCGCTGCTGCGGTACTG 240
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 241 GTTGCCAGTTCTATAAACAACGCGCATTTATCATCCACACATGCTTAGTGCGCTGCC 300
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 301 TATGGCTCGCTGGAAGTGATGGCGATTATA 333

RESULT 5

D90891

LOCUS

D90891 19150 bp DNA linear BCT 29-MAY-1997

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (sites)

Yamamoto,Y., Aiba,H., Baba,T., Hayashi,K., Inada,T., Isono,K.,

Itoh,T., Kimura,S., Kitagawa,M., Makino,K., Miki,T., Mitsuhashi,N.,

Mizobuchi,K., Mori,H., Nakade,S., Nakamura,Y., Nashimoto,H.,

Oehlma,T., Oyama,S., Saito,N., Sampei,G., Satoh,Y.,

Sivasubram,S., Tagami,H., Takahashi,H., Takeda,J., Takemoto,K.,

Uehara,K., Wada,C., Yamagata,S. and Horiuchi,T.

Construction of a contiguous 874-kb sequence of the Escherichia

coli -K12 genome corresponding to 50.0-68.8 min on the linkage map

and analysis of its sequence features

DNA Res. 4 (2), 91-113 (1997)

JOURNAL

PUBMED

9205837

REFERENCE

AUTHORS

2 (sites)

Aiba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiuchi,T.,

Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K.,

Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M.,

Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H.,

Motomura,K., Nakamura,Y., Nashimoto,H., Nishio,Y., Oshio,T.,

Saito,N., Sampei,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,

Yamamoto,Y. and Yano,M.

The systematic sequencing of the Escherichia coli genome in Japan

Unpublished

JOURNAL

REFERENCE

3 (bases 1 to 19150)

Mori,H.

TFDKRYLEDFADRVWVALTLAQDETLALQLTDEMLSGRFQFPATPTFLNCGKQQRG
ELVSCFLLRIDNNESIGRAVNSALQLSKRGQVAFLLSNLREAGAPIKRIENQSSGV
IPVMKLEDAFSAQAGAGAYLHAHPDILRFLDTKRENADEKIRIKTSLIG
VVIDITPHLAKENAQALFSPYDVYKFPADVAISOHYDELVADERIRKKYLANA
RDFORLAEIOFESGYPYIMVEDTVNRANPIAGRNNSNLCSEILOVNSASEYDENLD
YTFGHDSICNLGSLNHTAHTWSDPDPARTVETAVRGCLTAVSDMSHISVPSIEAGNA
SHAIGLQNMUNHLYAREGIAYSFEALDFNLIFYAITWHAULTSMLLARERGETFA
GFKQSRQAQW

CDS

6272..6964
/note="similar to [PIR Accession Number S34271]"
/codon_start=1
/transl_table=11
/product="ribonucleoside-diphosphate reductase (EC
1.17.4.1)"
/protein_id="BAA16540.1"
/translation="MVRKKWISPISSIMSPGMHCYPRCWHANAVKPSGNSHAM
PSGEYFSQYLQGNQPKTAKVGLPFRSGITLPREMWAQLRDDVMRYGIYNQLQAV
PPTGSISYINHTSSIHPIVAKVEIRKEGTRGVYYPAPFMNTNENLALYQDAVEIGAE

Alignment Scores:

Pred. No.: 9.59e-49 Length: 19150
Score: 552.00 Matches: 111
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-073-293A-6 (1-111) x D90891 (1-19150)

Qy 1 MetSerTyrGluValLeuLeuLeuGlyLeuLeuValGlyValAlaAseTyrCysPheArg 20
Db 13817 ATGAGCTATGAGGTTCTGCTTGGGTTACTAGTTGGCGTGGCGAATTATTGCTTCGCG 13876
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 13877 TATTTCGCGTGGCGCTCGGTGGGTAAATGCCGCCCAACACGTCGGCGGTAGGT 13936
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAla 60
Db 13937 ATTTTGTCTGCACACCATGGCATCGCTCGATATGCGCTCTGCTGGTGTCTCTACCGCA 13996
Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 13997 CCAGAGTGATGCACGATACACGCGTTTCGTGCCACGCTGGTGGCTTCGCGGTACTG 14056
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 14057 GGTGCCAGTTTCTATAAAACACGACGATATATCCCAACACTGTCTAGTGGCGCTGCC 14116
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 14117 TATGGGCTCGCTGGAAAGTGATGGCGATTATA 14149

RESULT 6
U00096_27
WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000

U00096_14 1400001 1510000
U00096_15 1500001 1610000
U00096_16 1600001 1710000
U00096_17 1700001 1810000
U00096_18 1800001 1910000
U00096_19 1900001 2010000
U00096_20 2000001 2110000
U00096_21 2100001 2210000
U00096_22 2200001 2310000
U00096_23 2300001 2410000
U00096_24 2400001 2510000
U00096_25 2500001 2610000
U00096_26 2600001 2710000
U00096_27 2700001 2810000
U00096_28 2800001 2910000
U00096_29 2900001 3010000
U00096_30 3000001 3110000
U00096_31 3100001 3210000
U00096_32 3200001 3310000
U00096_33 3300001 3410000
U00096_34 3400001 3510000
U00096_35 3500001 3610000
U00096_36 3600001 3710000
U00096_37 3700001 3810000
U00096_38 3800001 3910000
U00096_39 3900001 4010000
U00096_40 4000001 4110000
U00096_41 4100001 4210000
U00096_42 4200001 4310000
U00096_43 4300001 4410000
U00096_44 4400001 4510000
U00096_45 4500001 4610000
U00096_46 4600001 4639675

Continuation (28 of 47) of U00096 from base 2700001 (U00096 Escherichia coli K-12 MG1665)

Alignment Scores:

Pred. No.: 6.47e-48 Length: 110000
Score: 552.00 Matches: 111
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-073-293A-6 (1-111) x U00096_27 (1-110000)

Qy 1 MetSerTyrGluValLeuLeuLeuGlyLeuLeuValGlyValAlaAseTyrCysPheArg 20
Db 108366 ATGAGCTATGAGGTTCTGCTGGGTAACTAGTTGGCGTGGCGAATTATTGCTTCGCG 108425
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 108426 TATTTCGCGTGGCGCTCGGTGGGTAAATGCCGCCCAACACGTCGGCGGTAGGT 108485
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAla 60
Db 108486 ATTTTGTCTGCACACCATGGCATCGCTCGATATGCGCTCTGCTGGTGTCTCTACCGCA 108545
Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 108546 CCAGAGTGATGCACGATACACGCGTTTCGTGCCACGCTGGTGGCTTCGCGGTACTG 108605
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 108606 GGTGCCAGTTTCTATAAAACACGACGATATATCCCAACACTGTCTAGTGGCGTGCC 108665
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 108666 TATGGGCTCGCTGGAAAGTGATGGCGATTATA 108698

RESULT 7

U00096_28

WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Qy 1 MetSerTyrGluValLeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArg 20
 |||||
 Db 84241 ATGAGCTATGAGGTTCTGCTTGGGTACTGGTTGGCGGGCGAATTATTGCTTCGCG 84300
 |||||
 Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
 |||||
 Db 84301 TATTGGCGCTCGCGCTCGGTGGGTAAATGCCCGCCCAACCAACGTCGCGGTAGGT 84360
 |||||
 Qy 41 IleLeuLeuAspThrIleGlyLeAlaSerIleCysAlaLeuLeuValValSerThrAla 60
 |||||
 Db 84361 ATTTTGCTCGACACATTGGCATCGCTCGATATGGCTCTGCTGGTTGCTCTACCGCA 84420
 |||||
 Qy 61 ProGluValMetHisAspThrArgPheValProThrLeuValGlyPheAlaValLeu 80
 |||||
 Db 84421 CCAGAAGTGATGCGCATACACGCGTTCGTGGCCACGCTGGTTCGCGGTACTG 84480
 |||||
 Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
 |||||
 Db 84481 GGTGCCAGTTTCTATAAACACGCGAGCATTTATCATCCCAACACTGCTTAGTGGCTGGCC 84540
 |||||
 Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
 |||||
 Db 84541 TATGGGCTCGCTGGAAAGTGATGGCGATTATA 84573
 |||||

RESULT 9
 BA000007_35
 WPCOMMENT

Sequence split into 55 fragments LOCUS BA000007 Accession BA000007

Fragment Name	Begin	End
BA000007_00	1	110000
BA000007_01	100001	210000
BA000007_02	200001	310000
BA000007_03	300001	410000
BA000007_04	400001	510000
BA000007_05	500001	610000
BA000007_06	600001	710000
BA000007_07	700001	810000
BA000007_08	800001	910000
BA000007_09	900001	1010000
BA000007_10	1000001	1110000
BA000007_11	1100001	1210000
BA000007_12	1200001	1310000
BA000007_13	1300001	1410000
BA000007_14	1400001	1510000
BA000007_15	1500001	1610000
BA000007_16	1600001	1710000
BA000007_17	1700001	1810000
BA000007_18	1800001	1910000
BA000007_19	1900001	2010000
BA000007_20	2000001	2110000
BA000007_21	2100001	2210000
BA000007_22	2200001	2310000
BA000007_23	2300001	2410000
BA000007_24	2400001	2510000
BA000007_25	2500001	2610000
BA000007_26	2600001	2710000
BA000007_27	2700001	2810000
BA000007_28	2800001	2910000
BA000007_29	2900001	3010000
BA000007_30	3000001	3110000
BA000007_31	3100001	3210000
BA000007_32	3200001	3310000
BA000007_33	3300001	3410000
BA000007_34	3400001	3510000
BA000007_35	3500001	3610000
BA000007_36	3600001	3710000
BA000007_37	3700001	3810000
BA000007_38	3800001	3910000
BA000007_39	3900001	4010000
BA000007_40	4000001	4110000
BA000007_41	4100001	4210000
BA000007_42	4200001	4310000

BA000007_43 4300001 4410000
 BA000007_44 4400001 4510000
 BA000007_45 4500001 4610000
 BA000007_46 4600001 4710000
 BA000007_47 4700001 4810000
 BA000007_48 4800001 4910000
 BA000007_49 4900001 5010000
 BA000007_50 5000001 5110000
 BA000007_51 5100001 5210000
 BA000007_52 5200001 5310000
 BA000007_53 5300001 5410000
 BA000007_54 5400001 5498450

Continuation (36 of 55) of BA000007 from base 3500001 (BA000007 Escherichia coli O157:1)

Alignment Scores:
 Pred. No.: 1-75e-47 Length: 110000
 Score: 548.00 Matches: 110
 Percent Similarity: 99.1% Conservative: 0
 Best Local Similarity: 99.1% Mismatches: 1
 Query Match: 99.3% Indels: 0
 DB: 1 Gaps: 0

US-10-073-293A-6 (1-111) x BA000007_35 (1-110000)

Qy 1 MetSerTyrGluValLeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArg 20
 |||||
 Db 34502 ATGAGCTATGAGGTTCTGCTTGGGTACTGGTTGGCGGGCGAATTATTGCTTCGCG 34561
 |||||
 Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
 |||||
 Db 34562 TATTGGCGCTCGCGCTCGGTGGGTAAATGCCCGCCCAACCAACGTCGCGGTAGGT 34621
 |||||
 Qy 41 IleLeuLeuAspThrIleGlyLeAlaSerIleCysAlaLeuLeuValValSerThrAla 60
 |||||
 Db 34622 ATTTTGCTCGACACCATTTGGCATCGCTCGATATGGCTCTGCTGGTTGCTCTACCGCA 34681
 |||||
 Qy 61 ProGluValMetHisAspThrArgPheValProThrLeuValGlyPheAlaValLeu 80
 |||||
 Db 34682 CCAGAAGTGATGCGCATACACGCGTTCGTGGCCACGCTGGTTCGCGGTACTG 34741
 |||||
 Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
 |||||
 Db 34742 GGTGCCAGTTTCTATAAACACGCGCATTTATCATCCCAACACTGCTTAGTGGCTGGCC 34801
 |||||
 Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
 |||||
 Db 34802 TATGGGCTCGCTGGAAAGTGATGGCGATTATA 34834
 |||||

RESULT 10

AE016987

LOCUS Shigella flexneri 2a str. 2457T DNA linear BCT 22-APR-2003
 DEFINITION Shigella flexneri 2a str. 2457T section 10 of 16 of the complete genome.

ACCESSION AE016987 AE014073

VERSION AE016987.1 GI:30042162

KEYWORDS

SOURCE

ORGANISM

Shigella flexneri 2a str. 2457T

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Shigella.

REFERENCE

AUTHORS

Wei, J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W.,

Fournier, G., Mayhew, G.F., Plunkett, G. III, Rose, D.J., Darling, A.,

Mau, B., Perna, N.T., Payne, S.M., Runyen-Janecky, L.J., Zhou, S.,

Complete Genome Sequence and Comparative Genomics of Shigella

flexneri Serotype 2a Strain 2457T

Infect. Immun. 71 (5), 2775-2786 (2003)

JOURNAL

PUBMED

2 (bases 1 to 290380)

REFERENCE

AUTHORS

Wei, J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W.,

Fournier, G., Mayhew, G.F., Plunkett, G. III, Rose, D.J., Darling, A.,

Mau, B., Perna, N.T., Payne, S.M., Runyen-Janecky, L.J., Zhou, S.,

TITLE JOURNAL FEATURES source	Schwartz, D.C. and Blattner, F.R. Direct Submission Submitted (13-JUN-2002) Genetics Laboratory, University of Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA Location/Qualifiers 1. .290380 /organism="Shigella flexneri 2a str. 2457T" /mol_type="genomic DNA" /strain="2457T" /serotype="2a" /db_xref="taxon:198215" complement (82. .1209) /gene="gcpE" /locus_tag="S2733" complement (82. .1209) /gene="gcpE" /locus_tag="S2733" /note="residues 1 to 372 of 375 are 96.23 pct identical to residues 1 to 372 of 372 from Escherichia coli K-12 : B2515" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAP17888.1" /db_xref="GI:30042163" /translation="MKNQAPIQRKSTRIYGVNPIGGGAPIAVQSVMTNRTTDEATVNCIKALERVGADIVRVSVPTMDAAEAPFKLIKRVNVPVADIHFDYRIALKVAEYGVDCLRINPGENIERIRMVDCARDKNIPIRIGVNAGSLKDLQEKYGEPTQALLS AMRHVDRLNFPDQPKSVKASDVFLAVESYRIILAKIQIDQPLHLGITAGARGAVKSAIGLGLLSGGITLRLVSLAADPVVEIKVGFILKLSIRSRGINFTACPTCSRQ EFDVIGTVNALQRLEDITIPMDVSIIGCVVNGPEALVSTLGVGTGNGKKSGLYEDGV RKDRLNNDMDIQLEAIRAKASQLDEARRIDVQVGKLIIT" complement (1236. .2249) /gene="yfgA" /locus_tag="S2734" complement (1236. .2249) /gene="yfgA" /locus_tag="S2734" /function="putative membrane; Not classified" /note="residues 1 to 337 of 337 are 80.71 pct identical to residues 1 to 337 of 337 from Escherichia coli K-12 : B2516" /codon_start=1 /transl_table=11 /product="putative membrane protein" /protein_id="AAP17889.1" /db_xref="GI:30042164" /translation="MNTKATHDQNEALTTGABRLNAREQLGLSQQVAFERLCLKVSTV RDIEDKAPADLASTFLRGLYIRSVARLVHIPEELLPGLEKQAPLRAAKVAPMQSFSLGRRKKRDQWLMTFTVLVYVIGLSGAWQDHKAQOEIITWADQSSAEIJSNSEQ GQSVPLNTSTTIDPAPTSPASVDTTATNTQTPTAVTAPAVDPQONAVVSPQNAV DTAATPVTAAPTPDGAAPLTDQAGVTPPAADFNALVMNFTADQCLEVTDATGKKLF SGMQRKDGNNLTGQAPYKLGIPAAVQIQYQGPVDLSRFIRTNQVARTLNAEQSPAQ" complement (2534. .3688) /gene="yfgB" /locus_tag="S2735" complement (2534. .3688) /gene="yfgB" /locus_tag="S2735" /note="residues 1 to 384 of 384 are 99.47 pct identical to residues 1 to 384 of 384 from Escherichia coli K-12 : B2517" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAP17890.1" /db_xref="GI:30042165" /translation="MSEQLVTPENVTTKDGKINLLDLNRQMRBEFFKDLGKTFRADQ VMKWMHYCCDNFDMTDINKVLGKLEVAEIRAPEVVEQRSSDGTGIKWAIVGDQ RVETVYIPEDDRATLCVSSQVGCALCEKFCSTAQQGFNRNLRVSEIIIGQVWRAAKIV AAKVTGQRPIQNVNMGGEPLLNINVPVAMEIMLDDFGFLSKRRVTLSTSGVPPA LDKLGDMDVALAISLHAPNDEIRDSIVPINKKYNIEITFLAAVRRYLEKSNANQGRVT		
	gene		
	CDS		
	gene		
	CDS		
	gene		
	CDS		
	gene		
	CDS		
	gene		
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
CDS			
gene			
C			

repeat_region
LPTIKVPKRAPICSIKENEAQPNLLIGIEADGDIEIHIOATGSVATDTLPGDEPID
10137. .10170
/note="REP (repetitive extragenic palindromic) element,
contains 1 REP sequence"
complement(10179. .11462)
/gene="pep8"
/locus_tag="S2742"
complement(10179. .11462)
/gene="pep8"
/locus_tag="S2742"
/function="putative enzyme; Not classified"
/note="residues 1 to 427 of 427 are 92.50 pct identical to
residues 30 to 456 from Escherichia coli K-12 :
B2523"
/codon_start=1
/transl_table=1
/product="putative peptidase"
/protein_id="AAP17894.1"
/db_xref="GI:30042169"
/translation="MTSAMKITLSQPADARWGEKATYSINNDGITLHNGADDGLI
QRAKIDGLIGIKHVLSGEGWDADRCWAFQGYKPKGRKVEWPLDADAORDELIN
RLMIDVRDTINAPABELGPSOLAQRAVDLISNVAGDRVYRTTKGEDLRQGYMGL
HTVGRGSRSPVLLADYNPTGDKEAPVYACLVCKGKITPDSGGYSIKQTAPMSKSD
MGNAATVGAFAIMRGLNKRKVKFLCCADNLISGNAFLKGDIIITRNKGKKEVMNT
DAEGRVLADGLIDASAQKPEMIIDAATLTGAATGALNDYHALFSPDDALAGRLLAS
ASQNEPFWLPLAEFHRSQLPSNFASLNTNTGSAAYPAGASATAGFLSHFVYQQGW
LHIDCSATYRKAPVEQWSAGATGLGVRTIANLLTA"
complement(11521. .11721)
/gene="yfbJ"
/locus_tag="S2743"

Alignment Scores:

Pred. No.: 5,06e-47 Length: 290380
Score: 548.00 Matches: 110
Percent Similarity: 99.1% Conservativeness: 1
Best Local Similarity: 99.1% Mismatches: 0
Query Match: 99.3% Indels: 0
DB: 1 Gaps: 0

US-10-073-293a-6 (1-111) x AE016987 (1-290380)

Qy 1 MetSerTyrGluValLeuLeuLeuGlyLeuValGlyValAlaAenTyrCysPheArg 20
Db 155240 ATGAGCTATGAGGTTCTGCTGCTGGGTTACTGTTGGCGCGCAATTATTGCTTCGCG 155299
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 155300 TATTGGCGCTCGCGCTCGGTTGGGTAATGCCCGCCCAACCAACGTCGGCGGTAGGT 155359
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAla 60
Db 155360 ATTTGCTCGACACCATTTGCGATCGCTCGATATGCGCTCTGCTGGTTGCTCTACCGCA 155419
Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 155420 CCAGAACTGATGACAGTACACGCGCTTCTGTCGCCACGCTGCTGCGGTACTG 155479
Qy 81 GlyAlaSerPheThrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 155480 GGTGCCAGTTCTTATATAAACACGACGATTTATATCCCACTGCTTAGTCGCTGGCC 155539
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 155540 TATGGGCTCGCTGGAAAGTGATGGCGATTATA 155572

RESULT 11

AE016765
LOCUS
DEFINITION Escherichia coli CFT073 section 11 of 18 of the complete genome.
ACCESSION AE016765 AE014075
VERSION AE016765.1 GI:26109395

KEYWORDS

SOURCE
ORGANISM

Escherichia coli CFT073
Escherichia coli CFT073
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE
AUTHORS

1 (bases 1 to 305325)
Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,P.R.
Extensive Mosaic Structure Revealed by the Complete Genome Sequence
of Uropathogenic Escherichia coli
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 17020-17024 (2002)

JOURNAL
PUBMED

2 (bases 1 to 305325)
Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,P.R.
Direct Submission
Submitted (20-JUN-2002) Genetics Laboratory, University of
Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA

TITLE
JOURNAL

FEATURES
source
1. 305325
/organism="Escherichia coli CFT073"
/mol_type="genomic DNA"
/strain="CFT073"
/db_xref="taxon:199310"
complement(145. .438)
/locus_tag="c3147"
complement(145. .438)
/locus_tag="c3147"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="AAN81599.1"
/db_xref="GI:26109396"
/translation="MDITPFLHALCAVAQLVLFTGNWAYGAIAGCTFFIAREHTQ
AEYRWIEMFGHGKRIINPMWGGDFPRVWDVASLMDFAVPVVCALLVLLVNRG"
complement(481. .1521)
/locus_tag="c3148"
complement(481. .1521)
/locus_tag="c3148"
/notes="Residues 110 to 233 of 346 are 31.25 pct identical
to residues 486 to 611 of 669 from GenPept.129 :
>gb|AAD03286.1| (AF063097) gph [Enterobacteria phage P2]"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="AAN81600.1"
/db_xref="GI:26109397"
/translation="MATSTVTPDDIKTKSDVSKLNDQGSYATKLVDYSKDEIVGDM
SASWYQVLPSTGSAIFGRKLRSHTAGVEDAYCELYLKKWIDSPGNAMARLNDNG
TNICWDFTLNYGSTMIFPGDSYGLKMGNCNLSYSKRGSNALIKFDYTDITDQIKYANHG
STMLTNQTGTAAGVTAQLMGNSRPVYEVGVGGAYMFYAQNTDNTYMLSYNGAC
HATAFNQSDRLDKNIQVIDNRKNGYTYTLKENGMPYAGVIAQEALEAIP
VVGSAKYQDQAGSGSEGEGEREYTYDYSVGTGLLVQVARESDDRITALEEENALRQ
RLSAIAALASK"
complement(1531. .1812)
/locus_tag="c3149"
complement(1531. .1812)
/locus_tag="c3149"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="AAN81601.1"
/db_xref="GI:26109398"
/translation="MKDLTKFADRADFSAFMESIGYDDSMQGGDILIDVIGNVYKE
TGBLTEDGPVCVKEDGYFVNVRIINDAKKSSIFDKYAIIVVEHQLRGWM"
complement(1809. .4238)
/locus_tag="c3150"
complement(1809. .4238)
/locus_tag="c3150"
/notes="Residues 118 to 438 of 809 are 30.28 pct identical

Db 69012 CCAGAGGATGACGACGCGTTTCGTGCGCCACACTGGTGGCGTTCGCGGTACTG 69071
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 69072 GGTCCAGTTTCTATAAACAACGACGACGATTATCATCCCAACACTGCTTAGTGGCGTGGCC 69131
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 69132 TATGGGCTCGCTGGAAAGTGATGGCGATTATA 69164

RESULT 12

AE005174_35

WPCOMMENT

Sequence split into 56 fragments LOCUS AE005174 Accession AE005174

Fragment Name	Begin	End
AE005174_00	1	110000
AE005174_01	100001	210000
AE005174_02	200001	310000
AE005174_03	300001	410000
AE005174_04	400001	510000
AE005174_05	500001	610000
AE005174_06	600001	710000
AE005174_07	700001	810000
AE005174_08	800001	910000
AE005174_09	900001	1010000
AE005174_10	1000001	1110000
AE005174_11	1100001	1210000
AE005174_12	1200001	1310000
AE005174_13	1300001	1410000
AE005174_14	1400001	1510000
AE005174_15	1500001	1610000
AE005174_16	1600001	1710000
AE005174_17	1700001	1810000
AE005174_18	1800001	1910000
AE005174_19	1900001	2010000
AE005174_20	2000001	2110000
AE005174_21	2100001	2210000
AE005174_22	2200001	2310000
AE005174_23	2300001	2410000
AE005174_24	2400001	2510000
AE005174_25	2500001	2610000
AE005174_26	2600001	2710000
AE005174_27	2700001	2810000
AE005174_28	2800001	2910000
AE005174_29	2900001	3010000
AE005174_30	3000001	3110000
AE005174_31	3100001	3210000
AE005174_32	3200001	3310000
AE005174_33	3300001	3410000
AE005174_34	3400001	3510000
AE005174_35	3500001	3610000
AE005174_36	3600001	3710000
AE005174_37	3700001	3810000
AE005174_38	3800001	3910000
AE005174_39	3900001	4010000
AE005174_40	4000001	4110000
AE005174_41	4100001	4210000
AE005174_42	4200001	4310000
AE005174_43	4300001	4410000
AE005174_44	4400001	4510000
AE005174_45	4500001	4610000
AE005174_46	4600001	4710000
AE005174_47	4700001	4810000
AE005174_48	4800001	4910000
AE005174_49	4900001	5010000
AE005174_50	5000001	5110000
AE005174_51	5100001	5210000
AE005174_52	5200001	5310000
AE005174_53	5300001	5410000
AE005174_54	5400001	5510000
AE005174_55	5500001	5528445

Continuation (36 of 56) of AE005174 from base 3500001 (AE005174 Escherichia coli O157:H7)

Alignment Scores:

Pred. No.: 4,74e-47 Length: 110000
Score: 544.00 Matches: 109
Percent Similarity: 98.2% Conservative: 0
Best Local Similarity: 98.2% Mismatches: 2
Query Match: 98.6% Indels: 0
DB: 1 Gaps: 0

US-10-073-293A-6 (1-111) x AE005174_35 (1-110000)

Qy 1 MetSerTyrGluValLeuLeuGlyLeuValGlyValAlaAenTyrCysPheArg 20
Db 101792 ATGAGCTATGARGTCTGCTGTGGGTACTGGTTGGCGCGGAATTATGTCTTCGC 101851
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyValAlaValGly 40
Db 101852 TATTTGGCGCTGGCGCTGCGTGTGGGTAAATGCCGCCCAACCAACGTCGGCGGTAGT 101911
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAla 60
Db 101912 ATTTTGCTCGACACCAATTGGCATCGCTCGATATGCGTCTGCTGGTGTCTCTACCGCA 101971
Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 101972 CCAGAAAGTGATGCACGATACACGCCGTTTCGTGCCACGCTGGTGGCGGTACTG 102031
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 102032 GTTGCCAGTTTCTATAAACAACGACGATATATCATCCCAACACTGCTTARTGCCGTGCC 102091
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 102092 TATGGGCTCGCTGGAAAGTGATGGCGATTATA 102124

RESULT 13

AE005174_36

WPCOMMENT

Sequence split into 56 fragments LOCUS AE005174 Accession AE005174

Fragment Name	Begin	End
AE005174_00	1	110000
AE005174_01	100001	210000
AE005174_02	200001	310000
AE005174_03	300001	410000
AE005174_04	400001	510000
AE005174_05	500001	610000
AE005174_06	600001	710000
AE005174_07	700001	810000
AE005174_08	800001	910000
AE005174_09	900001	1010000
AE005174_10	1000001	1110000
AE005174_11	1100001	1210000
AE005174_12	1200001	1310000
AE005174_13	1300001	1410000
AE005174_14	1400001	1510000
AE005174_15	1500001	1610000
AE005174_16	1600001	1710000
AE005174_17	1700001	1810000
AE005174_18	1800001	1910000
AE005174_19	1900001	2010000
AE005174_20	2000001	2110000
AE005174_21	2100001	2210000
AE005174_22	2200001	2310000
AE005174_23	2300001	2410000
AE005174_24	2400001	2510000
AE005174_25	2500001	2610000
AE005174_26	2600001	2710000
AE005174_27	2700001	2810000
AE005174_28	2800001	2910000
AE005174_29	2900001	3010000
AE005174_30	3000001	3110000
AE005174_31	3100001	3210000
AE005174_32	3200001	3310000

JOURNAL Patent: US 6610836-A 3047 26-AUG-2003;
Genome Therapeutics Corporation; Waltham, MA

FEATURES
source Location/Qualifiers
1..372
 ~~/organism="unknown"~~
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 1-7e-34 Length: 372
Score: 403.00 Matches: 77
Percent Similarity: 84.1% Conservatives: 13
Best Local Similarity: 72.0% Mismatches: 17
Query Match: 73.0% Indels: 0
DB: Gaps: 0

US-10-073-293A-6 (1-111) x AR386318 (1-372)

QY 1 MetSerTyrGluValLeuLeuLeuGlyLeuValGlyValAlaIleTyrCysPheArg 20
 |||||
Db ATGAACAGTGAAGTCCTCTGTGGGGTGATAGTCGGCGGCGCAACTATCTCTTTCCG 96
QY 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaAargProThrIyaeRgLyAlaValGly 40
 |||||
Db TATCTTCCTCTGGCTTTGGCGCGGCCACCGTCGCCTGCGCGCGCGCGCGGTGAGC 156
QY 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAla 60
 ||:::
Db GTGCTGCTGATACTATTGGCATCGCTCGATTGGCGCTGCTGGTGGTCTCCAGCGTG 216
QY 61 ProGluValMetHisaspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
 ||:::
Db CCGGAGATCCTCGCCGATGCCATCGCTCGGCGCCGACGCTGACCGGCTTCTCTGGTCTG 276
QY 81 GlyAlaSerPheTyrllysThrArgSerIlellelleProThrIleuLeuSerAlaLeuAla 100
 ||:::
Db GGCGCGCGCTTCTGGAAACC CGCACATCATTTCCCCGACCTGCTGCGCGGTTTGCC 336
QY 101 TyrGlyLeuAlatrpLysVal 107
 ||:::
Db TATGGTCTGGCATGGAAATC 357

RESULT 15
ECH410307 3619 bp DNA linear BCT 15-APR-2005
LOCUS Erwinia chrysanthemi ygaZ gene (partial), ygaH gene, emrR gene,
DEFINITION emrA gene and emrB gene (partial).
ACCESSION AJ410307 GI:14970541
VERSION emrA gene; EmrA protein; emrB gene; EmrB protein; emrR gene; EmrR
KEYWORDS protein; ygaH gene; YgaH protein; ygaZ gene; YgaZ protein.
SOURCE Erwinia chrysanthemi (Pectobacterium chrysanthemi)
ORGANISM Erwinia chrysanthemi
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales;
Enterobacteriaceae; Pectobacterium.

REFERENCE 1
AUTHORS Reverchon,S.
TITLE Characterization of the Erwinia chrysanthemi multidrug resistance pump EmrAB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3619)
AUTHORS Reverchon,S.
TITLE Direct Submission
JOURNAL Submitted (12-JUL-2001) Reverchon S., Unite de Microbiologie et
Genetique, Institut des Sciences Appliquées de Lyon, Batiment Louis
Pasteur, 11 Avenue Jean Capelle, F69621 Villeurbanne cedex, FRANCE

FEATURES
source Location/Qualifiers
1..3619
 ~~/organism="Erwinia chrysanthemi"~~
 /mol_type="genomic DNA"
 /strain="j9937"
 /db_xref="taxon:556"
 <1..746

gene

CDS

```

/gene="ygaz"
<1..746
/gene="ygaz"
/function="unknown"
/codon_start=3
/transl_table=11
/product="ygaz protein"
/protein_id="CAC44347.1"
/db_xref="GI:14970542"
/db_xref="InterPro:IPR011606"
/db_xref="UniProt/TREMBL:Q93KB4"
/translations="NSVSSQADARTTADASPAQLDPLPIVIGYVPVAFAGLNAY
KLGTPPEALFSLCIIYAGASQFVITALLSAGASIWVAALTYWAMDVRHVLVYGSLSR
RIGWRLPTKTAWAFGLTDEVFAAATLARLDRNRSEPMGLGVALSALWSVAGTV
IGAVFRNGLEGYPAVEAALAFMLPLFLSLASFRRLQSLVVAALGGACLGLLVS
SIPAAIIGIGGGCLASLYQPTSVKEETR"
743..1084
/gene="ygaH"
743..1084
/gene="ygaH"
/function="unknown"
/codon_start=1
/transl_table=11
/product="ygaH protein"
/protein_id="CAC44348.1"
/db_xref="GI:14970543"
/db_xref="UniProt/TREMBL:Q93KB3"
/translations="MNTSVLLIGLVGVNPLFRYLPLRLGAGRTSGLNLRGRRTALLL
DSIGIASICALLVSTLPDIMQHTKMLPLTAGFALLTLCFYKTRSVLSTLGCALCY
GIVFKWFOQWA"
1259..1771
/gene="emrR"
1259..1771
/gene="emrR"
/function="negative regulator"
/codon_start=1
/transl_table=11
/product="EmrR protein"
/protein_id="CAC44349.1"
/db_xref="GI:14970544"
/db_xref="GOA:Q93KB2"
/db_xref="InterPro:IPR000835"
/db_xref="InterPro:IPR011991"
/db_xref="UniProt/TREMBL:Q93KB2"
/translations="MESSFAPIEDMLMRASRPDPFYREVLLRLFLHMOTKLSHR
NRMLKEQDINETLPMALLTLESQENYCIOPSELSAALGSWRTNATRIADELEKRGWIE
RRSDSDRRLCYLYMTEKGAFLDELLPPQHRSLNLTCSALEDSESDQLEVLNRKLLQ
RLDEMDQDGI"
1946..3121
/gene="emrA"
1946..3121
/gene="emrA"
/function="membrane fusion protein"
/codon_start=1
/transl_table=11
/product="EmrA protein"
/protein_id="CAC44350.1"
/db_xref="GI:14970545"
/db_xref="GOA:Q93KB1"
/db_xref="InterPro:IPR005694"
/db_xref="InterPro:IPR006143"
/db_xref="UniProt/TREMBL:Q93KB1"
/translations="MSANVEKQTPPTWPSKSKSRKGIITLLMIVFFFIGVWFAYVY
LVLRHQETDDAVAGNQIQIMSVQSVARVNVNDTFVKGGDLVVELDPTDAEQAF
ERSKTTIANSVRVHQHMNVROIQANIDLQQLDAKDKAISDLNRREVLGRANAIGRED
LQHQARQASARASLEAKQOYATQALVLTSLQOQPAIAQAATDLRNALWALKEITH
IVSPVDGYSVRRVQVGRVTPSSALMAVVPAPLWDFANFKETQLANRIGQPATII
SDLYGDSVYKGVGLDMGTGSAFLPQANATGNWIKVVERLPVRIELDPKQLTDY
PLRIGLSTLVKVDNTDGTGKWLSDASRATPAYQSDALDLDLSPVNMQIDQIIGANAG"
3139..3619
/gene="emrB"
3139..3619
/gene="emrB"

```

gene

CDS

ORIGIN

```

Alignment Scores:
Pred. No.: 9,82e-24 Length: 3619
Score: 313.50 Matches: 63
Percent Similarity: 72.9% Conservative: 15
Best Local Similarity: 58.9% Mismatches: 28
Query Match: 56.8% Indels: 1
DB: 1 Gaps: 1

```

US-10-073-293A-6 (1-111) x ECH410307 (1-3619)

```

Qy 1 MetSerTyrGluValLeuLeuLeuGlyLeuValGlyValAlaLeuTyrCysPheArg 20
Db 743 ATGAATACTTCAGTATTATTGATTGGTTTACTGTGGGAACAGTTAACTTCTCTGTTCCGA 802
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThr---LysArgGlyAlaVal 39
Db 803 TACCTGCGCTGCGGTTGGCGCAGCGCGACGTCGGTAACCTGCGCGCGGCGAGAAC 862
Qy 40 GlyIleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThr 59
Db 863 GCCCTGCTGCTCGACACGATCGGCATTGGCTCGATTGTGCGCTGCTGTGTATCCACC 922
Qy 60 AlaProGluValMethIleAspThrArgArgPheValProThrLeuValGlyPheAlaVal 79
Db 923 CTCGCCGATATCATGCACGACACGCGAAAGATGTTACCGACGCTGGCGGGGTTTCGCCCTG 982
Qy 80 LeuGlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeu 99
Db 983 CTGACGCTTTGTTTTTACAAAACCCGACGATCGTACTATCAACGCTGCTGCGCGGCTG 1042
Qy 100 AlaTyrGlyLeuAlaTrpLys 106
Db 1043 TGCTACGGCATAGTATTATAA 1063

```

Search completed: February 16, 2006, 19:47:18
 Job time : 2895.6 secs

gene

CDS

```

/gene="emrA"
1946..3121
/gene="emrA"
/function="membrane fusion protein"
/codon_start=1
/transl_table=11
/product="EmrA protein"
/protein_id="CAC44350.1"
/db_xref="GI:14970545"
/db_xref="GOA:Q93KB1"
/db_xref="InterPro:IPR005694"
/db_xref="InterPro:IPR006143"
/db_xref="UniProt/TREMBL:Q93KB1"
/translations="MSANVEKQTPPTWPSKSKSRKGIITLLMIVFFFIGVWFAYVY
LVLRHQETDDAVAGNQIQIMSVQSVARVNVNDTFVKGGDLVVELDPTDAEQAF
ERSKTTIANSVRVHQHMNVROIQANIDLQQLDAKDKAISDLNRREVLGRANAIGRED
LQHQARQASARASLEAKQOYATQALVLTSLQOQPAIAQAATDLRNALWALKEITH
IVSPVDGYSVRRVQVGRVTPSSALMAVVPAPLWDFANFKETQLANRIGQPATII
SDLYGDSVYKGVGLDMGTGSAFLPQANATGNWIKVVERLPVRIELDPKQLTDY
PLRIGLSTLVKVDNTDGTGKWLSDASRATPAYQSDALDLDLSPVNMQIDQIIGANAG"
3139..3619
/gene="emrB"
3139..3619
/gene="emrB"

```

gene

CDS

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:00:29 ; Search time 333.935 Seconds
(without alignments)
2215.341 Million cell updates/sec

Title: US-10-073-293A-6
Perfect score: 552
Sequence: 1 MSVEVLGLLLGVANVCFR'.....IPTLLSALAYGLAWKVAII 111

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/abs/ABSSWEB spool/US10073293/runat_15022006_120733_951/app_query.fasta.1
-DB=N_Geneseq -QFMT=fascap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs07
-USR=US10073293 @CGN 1 1 900 @runat_15022006_120733_951 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*
14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	552	100.0	336	6	ABV75657 E. coli L
2	552	100.0	336	14	ADZ39952
3	552	100.0	336	14	ADZ36196
4	423	76.6	314	4	AAK70628

5	403	73.0	372	11	ACH97252
6	325	58.9	238	4	AAK57571
7	271.5	49.2	339	10	ACF69126
8	271.5	49.2	110000	10	ACF67367_20
9	271.5	49.2	243072	10	ACF65382
10	182.5	33.1	369	9	ADA31451
11	164	29.7	100	8	ACD80498
12	160	29.0	100	8	ACD80497
13	155	28.1	100	8	ACD80499
14	98	17.8	172	10	ACF79478
15	78	14.1	1593	8	ACA39729
16	78	14.1	1879	4	AAK505386
17	78	14.1	1879	6	ABA00331
18	77.5	14.0	1326	8	ACA38890
19	77.5	14.0	19988	4	AAF28523
20	77	13.9	324	12	ADK13752
21	77	13.9	330	4	AAH84429
22	77	13.9	333	3	AAK66035
23	77	13.9	333	4	AAK52620
24	77	13.9	333	4	AAH84506
25	77	13.9	333	5	AAH81328
26	77	13.9	333	8	ACA18608
27	77	13.9	333	8	ACA13378
28	77	13.9	924	13	ADS48786
29	77	13.9	10944	4	AAK46250
30	76	13.8	330	8	ACA49494
31	76	13.8	333	4	AAK56235
32	76	13.8	333	8	ACA19680
33	76	13.8	333	8	ACA51662
34	76	13.8	405	8	ACA49545
35	76	13.8	1544	6	ABX97164
36	76	13.8	1544	12	ADN62044
37	76	13.8	1780	3	AAK32850
38	76	13.8	1950	8	ACA46614
39	76	13.8	1968	6	ABN93231
40	76	13.8	1968	13	ADS04188
41	76	13.8	4041	4	AAH54209
42	76	13.8	4342	4	AAH54625
43	75.5	13.7	999	9	ADB11933
44	75.5	13.7	999	9	ADB11925
45	75.5	13.7	999	9	ADB11921

ALIGNMENTS

RESULT 1
ABV75657
ID ABV75657 standard; DNA; 336 BP.
XX
AC ABV75657;
XX
DT 29-JAN-2003 (first entry)
XX
DE E. coli L-amino acid producing gene #2.
XX
KW L-amino acid; gene; ss; E.coli.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT CDS 1..336
FT /*tag= a
XX
PN EP1239041-A2.
XX
PD 11-SEP-2002.
XX
PF 13-FEB-2002; 2002EP-00003335.
XX
PR 13-FEB-2001; 2001RU-00103865.
PR 26-FEB-2001; 2001RU-00104998.
PR 26-FEB-2001; 2001RU-00104999.
PR 28-JUN-2001; 2001RU-00117632.

```
PR 28-JUN-2001; 2001RU-00117633.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Gussyatiner MM;
XX WPI; 2002-699856/76.
XX P-PSDB; ABP57757.
XX
XX Novel L-amino acid producing Escherichia bacterium, is modified to
PT enhance L-amino acid production by enhancing the activities of protein
PT capable of making bacterium to have enhanced resistance to L-amino acids.
XX
XX Claim 1; Page 24; 33pp; English.
XX
XX The invention relates to a novel L-amino acid producing bacterium of
CC genus Escherichia, modified to enhance L-amino acid production by
CC enhancing the activities of a protein. The novel bacterium is useful for
CC producing L-amino acid e.g. L-Thr, L-Val, L-Pro, L-Leu, L-Met and L-Arg,
CC by cultivating the bacterium in a culture medium and collecting L-amino
CC acid to be produced and accumulated from the culture medium. The present
CC sequence represents a gene of the invention which causes increased L-
CC amino acid production in E. coli
XX
XX Sequence 336 BP; 62 A; 90 C; 93 G; 91 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.01e-65 Length: 336
Score: 552.00 Matches: 111
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-073-293A-6 (1-111) x ABV75657 (1-336)
QY 1 MetSerTyrGluValLeuLeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArg 20
DB 1 ATGAGCTATGAGTTCTGCTGTTGGGTACTAGTGGCGTGGCGAATTATTGCTTCGCG 60
QY 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyValAlaValGly 40
DB 61 TATTGGCGCTGGCGTGGTGGTAAATGCCGCCACCAACGTCGGCGCGGTAGGT 120
QY 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAla 60
DB 121 ATTTTGTCTGCACACCATTTGGCATGCGCTCGATATGCGCTCTGCTGGTGTCTCTACCGCA 180
QY 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
DB 181 CCAGAAGTGATGCACGATACACGCGTTTCGTGCCACGCTGGTGGCTTCCGCGGTACTG 240
QY 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
DB 241 GGTGCCAGTTCTTATATAAACACGACGAGCATTTATCATCCACACACTGCTTAGTGGCGTGGCC 300
QY 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
DB 301 TATGGGCTCGCTCGAAAGTGATGGCGATTATA 333
RESULT 2
ADZ39952
ID ADZ39952 standard; cDNA; 336 BP.
XX
XX AC ADZ39952;
XX
XX DT 14-JUL-2005 (first entry)
XX
XX DE E.coli ygaH gene for improved amino acid production in bacteria.
XX
XX KW amino acid production; fermentation; transport protein; gene; ss.
XX
XX OS Escherichia coli.
XX
```

```
XX
XX Key Location/Qualifiers
XX CDS 1..336
XX /tag= a
XX /product= "amino acid biosynthesis protein"
XX /function= "putative transport protein"
XX /gene= "ygaH"
XX
XX EP1526179-A1.
XX
XX 27-APR-2005.
XX
XX 13-FEB-2002; 2004EP-00028876.
XX
XX 13-FEB-2001; 2001RU-00103865.
XX 26-FEB-2001; 2001RU-00104998.
XX 26-FEB-2001; 2001RU-00104999.
XX 28-JUN-2001; 2001RU-00117632.
XX 28-JUN-2001; 2001RU-00117633.
XX 13-FEB-2002; 2002EP-00003335.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Gussyatiner MM;
XX WPI; 2005-308098/32.
XX P-PSDB; ADZ39953.
XX GENBANK; U00096.
XX
XX New L-amino acid producing bacterium belonging to the genus Escherichia,
XX useful for producing L-amino acids, e.g. L-threonine, L-valine, L-
XX proline, L-methionine, or L-arginine.
XX
XX Disclosure; SEQ ID NO 5; 35pp; English.
XX
XX The invention relates to an L-amino acid producing bacterium belonging to
XX the genus Escherichia, where the bacterium has been modified so that the
XX L-amino acid production by the bacterium is enhanced by enhancing
XX activities of proteins by transformation of the bacterium with DNA coding
XX for protein or by alteration of an expression regulation sequence of the
XX DNA on the chromosome of the bacterium. The L-amino acid producing
XX bacterium is useful for producing L-amino acids by fermentation,
XX including L-threonine, L-valine, L-proline, L-methionine, or L-arginine.
XX The genes are useful for improving L-amino acid productivity. This
XX sequence corresponds to the ygaH gene encoding protein b2683, one of the
XX L-amino acid biosynthesis pathway proteins. The protein is a putative
XX transport protein within the biosynthesis pathway. Enhancement of the
XX expression of this gene increases production on L-threonine, L-valine, L-
XX leucine, L-proline and L-methionine.
XX
XX Sequence 336 BP; 62 A; 90 C; 93 G; 91 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.01e-65 Length: 336
Score: 552.00 Matches: 111
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0
US-10-073-293A-6 (1-111) x ADZ39952 (1-336)
QY 1 MetSerTyrGluValLeuLeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArg 20
DB 1 ATGAGCTATGAGTTCTGCTGTTGGGTACTAGTGGCGTGGCGAATTATTGCTTCGCG 60
QY 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyValAlaValGly 40
DB 61 TATTGGCGCTGGCGCTGGTGGTAAATGCCGCCACCAACGTCGGCGCGGTAGGT 120
QY 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAla 60
DB 121 ATTTTGTCTGCACACCATTTGGCATGCGCTCGATATGCGCTCTGCTGGTGTCTCTACCGCA 180
DB 121 ATTTTGTCTGCACACCATTTGGCATGCGCTCGATATGCGCTCTGCTGGTGTCTCTACCGCA 180
```


Qy 61 ProGluValMetHisAspThrArgPheValProThrLeuValGlyPheAlaValLeu 80
 |||||
 Db 181 CCAGAAGTGATGCACGATACACGCGTTTCGTGCCACGCTGTCGCGTTCGCGGTACTG 240
 |||||
 Qy 81 GlyAlaSerPheThrArgSerIleIleProThrLeuLeuSerAlaLeuAla 100
 |||||
 Db 241 GGTGCCAGTTTCTATAAACACGACGATTCATCCACACACTGCTTAGTGGCGCTGGCC 300
 |||||
 Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
 |||||
 Db 301 TATGGCTCGCTGGAAGTGATGGCGATTATA 333
 |||||

RESULT 3

ADZ36196

ID ADZ36196 standard; DNA; 336 BP.

XX AC

ADZ36196;

XX DT 14-JUL-2005 (first entry)

XX DE Escherichia coli b2683 encoding DNA SEQ ID NO:5.

XX KW amino acid production; gene; ds.

XX OS Escherichia coli.

XX PH Key Location/Qualifiers

XX FT 1..336

XX FT /*tag= a

XX FT /product= "b2683 protein"

XX PN EP1526181-A1.

XX PD 27-APR-2005.

XX PF 13-FEB-2002; 2004EP-00028877.

XX PR 13-FEB-2001; 2001RU-00103865.

XX PR 26-FEB-2001; 2001RU-00104998.

XX PR 26-FEB-2001; 2001RU-00104999.

XX PR 28-JUN-2001; 2001RU-00117632.

XX PR 28-JUN-2001; 2001RU-00117633.

XX PR 13-FEB-2002; 2002EP-00003335.

XX PA (AJIN) AJINOMOTO CO INC.

XX PI Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Guseyatiner MM;

XX DR WPI; 2005-317157/33.

XX DR P-PSDB; ADZ36197.

XX PT New modified L-amino acid producing bacterium useful to enhance the

XX PT production of L-amino acid by enhancing the activities of proteins in a

XX PT cell of bacterium.

XX PS Disclosure; SEQ ID NO 5; 35pp; English.

XX CC The invention relates to an L-amino acid (A) producing bacterium (I)

XX CC (belonging to the genus Escherichia), which is modified to enhance the

XX CC production of (A) by enhancing the activities of proteins (G) or (H) in a

XX CC cell of (I). Also described is a method for producing (A) comprising

XX CC cultivating the bacterium in a culture medium and collecting the produced

XX CC and accumulated L-amino acid. The modified bacterium has the ability to

XX CC grow on a minimal medium containing L-amino acid or its analog in a

XX CC minimal concentration and ability to grow faster on a medium containing L

XX CC -amino acid or its analog than the unmodified strain or the wild type

XX CC strain, or the parental strain of the bacterium. The present sequence

XX CC encodes the E. coli b2683 protein which has L-amino acid excretion

XX CC activity.

XX SQ Sequence 336 BP; 62 A; 90 C; 93 G; 91 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.01e-65 Length: 336
 Score: 552.00 Matches: 111
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0

US-10-073-293A-6 (1-111) x ADZ36196 (1-336)

Qy 1 MetSerTyrGluValLeuLeuGlyLeuValGlyValAlaAenTyrCysPheArg 20
 |||||
 Db 1 ATGAGCTATGAGGTTCTGCTGCTGGGTTACTAGTTGGCGTGGCGAATTTATTTGCTTCGCG 60
 |||||
 Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
 |||||
 Db 61 TATTTGGCGTGGCGCTGCTGGTGGTAAAGCCGCCCAACCAACGTTGGCGCGGTAGGT 120
 |||||
 Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAla 60
 |||||
 Db 121 ATTTTGGCTCGACACCATTTGGCATCGCTCGATATGGCTCTGCTGGTTGTCTTACCGCA 180
 |||||
 Qy 61 ProGluValMethisaspThrArgPheValProThrLeuValGlyPheAlaValLeu 80
 |||||
 Db 181 CCAGAAAGTGATGCACGATACACGCGTTTCGTGCCACGCTGGTTCGCGGTACTG 240
 |||||
 Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
 |||||
 Db 241 GGTGCCAGTTTCTATAAACACGACGATTCATCCCAACACTGCTTAGTGGCGCTGGCC 300
 |||||
 Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
 |||||
 Db 301 TATGGCTCGCTGGAAAGTGATGGCGATTATA 333
 |||||

RESULT 4

AAK70628

ID AAK70628 standard; DNA; 314 BP.

XX AC

AAK70628;

XX DT 06-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25440.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX OS Homo sapiens.

XX PN WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001354.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180828P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225274P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0244647P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosure; SEQ ID NO 25440; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169

CC represent sequences used in the exemplification of the present invention
XX Sequence 314 BP; 69 A; 86 C; 74 G; 85 T; 0 U; 0 Other;
SQ

Alignment Scores: 3.74e-48 Length: 314
Pred. No.: 423.00 Matches: 86
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 76.6% Indels: 0
Query Match: 4 Gaps: 0
DB:

US-10-073-293A-6 (1-111) x AAK70628 (1-314)

Qy 26 LeuArgValGlyAanAlaAatgProThrLysArgGlyAlaValGlyIleLeuLeuAspThr 45
Db 3 CTGGGTGGGTAAATGCGCGCCCAACCAAGTCGGGGTAGTATTTGCTGCACAC 62
Qy 46 IleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAlaProGluValMetHis 65
Db 63 ATTGGCATCGCTCGATATGCGCTCTGCTGGTTGCTCTACCGCACCAAGTGTATGCAC 122
Qy 66 AspThrArgArgPheValProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyr 85
Db 123 GATACACGCCGTTTGGTGCCACACGCTGCTGGCTTCGGGTACTGGGTGCAGTTCTAT 182
Qy 86 LysThrArgSerIleIleProThrLeuLeuSerAlaLeuAlaTyrGlyLeuAlaTyr 105
Db 183 AAACACGCGCATTATATCAATCCCAACTGCTTAGTGGCTGGGCTATGGCTGGCTGG 242
Qy 106 LysValMetAlaIleIle 111
Db 243 AAAGTGATGCGATTATA 260

RESULT 5

ACH97252
ID ACH97252 standard; DNA; 372 BP.

XX AC ACH97252;

XX DT 29-JUL-2004 (first entry)

XX DE Klebsiella pneumoniae polynucleotide seqid 3047.

XX KW Recombinant expression vector; transcription regulatory element;

XX KW Klebsiella pneumoniae protein; antibacterial; vaccine; gene; ds.

XX OS Klebsiella pneumoniae.

XX PN US6610836-B1.

XX PD 26-AUG-2003.

XX PF 27-JAN-2000; 2000US-00489039.

XX PR 29-JAN-1999; 99US-0117747P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Breton GL, Osborne M;

XX DR WPI; 2003-895346/82.

XX DR P-PSDB; ABO63701.

XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.

XX PS Disclosure; SEQ ID NO 3047; 932pp; English.

XX CC The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression

CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention

XX SQ Sequence 372 BP; 50 A; 123 C; 113 G; 86 T; 0 U; 0 Other;

Alignment Scores: 2.55e-45 Length: 372
Pred. No.: 403.00 Matches: 77
Score: 84.1% Conservative: 13
Percent Similarity: 72.0% Mismatches: 17
Best Local Similarity: 73.0% Indels: 0
Query Match: 11 Gaps: 0
DB:

US-10-073-293A-6 (1-111) x ACH97252 (1-372)

Qy 1 MetSerTyrGluValLeuLeuLeuGlyLeuLeuValGlyValAlaAanTyrCysPheArg 20
Db 37 ATGAACAGTGAAGTCCCTTCTGCTCGGGGTGATAGTCGGCGCGCTCAACTATCTCTTTCG 96
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 97 TATCTTCTCTGCTGGTTGCGCGCCACGCTGCTCTGCCCGCCCGCGCGCTGAGC 156
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAla 60
Db 157 GTGCTGCTGATACTATTGGCATCGCTCGATTTGGCGCTGCTGTGGTGTCTCCAGCGTG 216
Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 217 CCGGAGATCTCTCGCGATGCCATCGCTGCGCGCCGACGCTGACCGGCTTCTCTGGTGTG 276
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 277 GGGGCGCCCTCTGGAAACCCCGCAGCATATTGTCCCGACCTGCTGTGGCGTTTGGCC 336
Qy 101 TyrGlyLeuAlaTyrLysVal 107
Db 337 TATGGTCTGGCATGGAAATC 357

RESULT 6

AAK57571

ID AAK57571 standard; cDNA; 238 BP.

XX AC AAK57571;

XX DT 06-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:2631.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.

XX OS Homo sapiens.

XX PN WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001354.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184564P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

PR	07-JUL-2000;	2000US-0216880P.	PR	20-OCT-2000;	2000US-0241809P.
PR	11-JUL-2000;	2000US-0217487P.	PR	20-OCT-2000;	2000US-0241826P.
PR	11-JUL-2000;	2000US-0217496P.	PR	01-NOV-2000;	2000US-0244617P.
PR	14-JUL-2000;	2000US-0218290P.	PR	08-NOV-2000;	2000US-0246474P.
PR	26-JUL-2000;	2000US-0220963P.	PR	08-NOV-2000;	2000US-0246475P.
PR	26-JUL-2000;	2000US-0220964P.	PR	08-NOV-2000;	2000US-0246476P.
PR	14-AUG-2000;	2000US-0224518P.	PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246611P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246613P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-02266681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-02266868P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	11-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	05-JAN-2001;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR		
PR	25-SEP-2000;	2000US-0234998P.	PR		
PR	26-SEP-2000;	2000US-0235484P.	PR		
PR	27-SEP-2000;	2000US-0235834P.	PR		
PR	27-SEP-2000;	2000US-0235836P.	PR		
PR	29-SEP-2000;	2000US-0236327P.	PR		
PR	29-SEP-2000;	2000US-0236367P.	PR		
PR	29-SEP-2000;	2000US-0236368P.	PR		
PR	29-SEP-2000;	2000US-0236369P.	PR		
PR	29-SEP-2000;	2000US-0236370P.	PR		
PR	02-OCT-2000;	2000US-0236802P.	PR		
PR	02-OCT-2000;	2000US-0237037P.	PR		
PR	02-OCT-2000;	2000US-0237038P.	PR		
PR	02-OCT-2000;	2000US-0237039P.	PR		
PR	13-OCT-2000;	2000US-0237040P.	PR		
PR	13-OCT-2000;	2000US-0239933P.	PR		
PR	13-OCT-2000;	2000US-0239937P.	PR		
PR	20-OCT-2000;	2000US-0240960P.	PR		
PR	20-OCT-2000;	2000US-0241221P.	PR		
PR	20-OCT-2000;	2000US-0241785P.	PR		
PR	20-OCT-2000;	2000US-0241786P.	PR		
PR	20-OCT-2000;	2000US-0241787P.	PR		
PR	20-OCT-2000;	2000US-0241808P.	PR		

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

P-PSDB; AAM84790.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Claim 1; SEQ ID NO 2631; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I)

CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent the
 CC diagnosis and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 238 BP; 43 A; 71 C; 61 G; 60 T; 0 U; 3 Other;

Alignment Scores:
 Pred. No.: 6.03e-35 Length: 238
 Score: 325.00 Matches: 73
 Percent Similarity: 93.7% Conservative: 1
 Best Local Similarity: 92.4% Mismatches: 4
 Query Match: 58.9% Indels: 2
 DB: 4 Gaps: 0

US-10-073-293A-6 (1-111) x AAK57571 (1-238)

Qy 26 LeuArgValGlyAenAlaArgProThrLysArgGlyAlaValGlyIleLeuLeuAspThr 45
 Db 3 CTGGGTGGTAAATGCGCCGCCAACCAAGTGGCGGTAGTATTTCTCGACACC 62
 Qy 46 IleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAlaProGluValMetHis 65
 Db 63 ATTGGCATCG-TGATATGCGTCTGCTGTTGCTCTACCGCACCAAGATGATGAC 121
 Qy 66 AspThrArgArgPheValProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyr 85
 Db 122 GATACACGCGTTTCGTCGCCACGCTGCTGCGCTTCCGCTACTGCGTGCACATTTCTAT 181
 Qy 86 LysThrArgSerIleIleProThrLeuLeuSerAla-LeuAlaTyrGlyLeu 103
 Db 182 AAAACAGCAGCATATCATCCACACTGGTAGTGGCTTGGCCCTAAGGGCTT 236

RESULT 7
 ACF69126
 ID ACF69126 standard; DNA; 339 BP.
 XX
 AC ACF69126;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens nucleotide sequence #7593.
 XX

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 XX whooping cough; gene; ds.
 XX
 OS Photorhabdus luminescens.
 XX
 PN WO200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX

XX (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PA
 XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunat F, Danchin A;
 PI Buchrieser C;
 XX
 DR WPI; 2003-148459/14.
 XX
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX

PS Claim 2; SEQ ID NO 7593; 1205pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly of human diseases for which P.
 CC sequence represents one of the isolated P. luminescens genes
 XX
 SQ Sequence 339 BP; 99 A; 66 C; 62 G; 112 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2e-27 Length: 339
 Score: 271.50 Matches: 54
 Percent Similarity: 72.0% Conservative: 23
 Best Local Similarity: 50.5% Mismatches: 25
 Query Match: 49.2% Indels: 5
 DB: 10 Gaps: 2

US-10-073-293A-6 (1-111) x ACF69126 (1-339)

Qy 4 GluValLeuLeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArgTyrLeuPro 23
 Db 13 AAGATTTTGGTATGGAGCTATTGTTGGGTAGCTAACTTTTCATTTCTGCTATCTGCCA 72
 Qy 24 LeuArgLeuArgValGlyAsnAlaArgProThr-----LysArgGlyAlaValGly 40
 Db 73 CTACGATTT-----GGGAAAGCAGCAATCTGCCGCGAGAAAGCTGGAAACAAAGC 126
 Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAla 60
 Db 127 ATTATCTTGACAGTATGGTATGTCATTCATTTGTTTCTTCTATCATCGTATCAGGTGTA 186
 Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
 Db 187 CCTGATGTGATGAGAGAAAGTCAAAACTACTTCTTACCTCATAGGTTGCTGACCATC 246
 Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleProThrLeuLeuSerAlaLeuAla 100
 Db 247 TGTTTAGTCTTTTACAAAACAAAGCAAAATTTATCTCGCAACACTATTTTGGCGCACTGCT 306
 Qy 101 TyrGlyLeuAlaTrpLysVal 107
 Db 307 TTTGGACTAATCAATCAATA 327

RESULT 8

ACF67367 20
 Continuation (21 of 57) of ACF67367 from base 2000001 (Photorhabdus luminescens nucleot
 WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367
 WP Fragment Name Begin End
 WP ACF67367_00 1 110000
 WP ACF67367_01 100001 210000
 WP ACF67367_02 200001 310000
 WP ACF67367_03 300001 410000
 WP ACF67367_04 400001 510000
 WP ACF67367_05 500001 610000
 WP ACF67367_06 600001 710000
 WP ACF67367_07 700001 810000

WP	ACF67367_08	800001	910000	
WP	ACF67367_09	900001	1010000	
WP	ACF67367_10	1000001	1110000	
WP	ACF67367_11	1100001	1210000	
WP	ACF67367_12	1200001	1310000	
WP	ACF67367_13	1300001	1410000	
WP	ACF67367_14	1400001	1510000	
WP	ACF67367_15	1500001	1610000	
WP	ACF67367_16	1600001	1710000	
WP	ACF67367_17	1700001	1810000	
WP	ACF67367_18	1800001	1910000	
WP	ACF67367_19	1900001	2010000	
WP	ACF67367_20	2000001	2110000	
WP	ACF67367_21	2100001	2210000	
WP	ACF67367_22	2200001	2310000	
WP	ACF67367_23	2300001	2410000	
WP	ACF67367_24	2400001	2510000	
WP	ACF67367_25	2500001	2610000	
WP	ACF67367_26	2600001	2710000	
WP	ACF67367_27	2700001	2810000	
WP	ACF67367_28	2800001	2910000	
WP	ACF67367_29	2900001	3010000	
WP	ACF67367_30	3000001	3110000	
WP	ACF67367_31	3100001	3210000	
WP	ACF67367_32	3200001	3310000	
WP	ACF67367_33	3300001	3410000	
WP	ACF67367_34	3400001	3510000	
WP	ACF67367_35	3500001	3610000	
WP	ACF67367_36	3600001	3710000	
WP	ACF67367_37	3700001	3810000	
WP	ACF67367_38	3800001	3910000	
WP	ACF67367_39	3900001	4010000	
WP	ACF67367_40	4000001	4110000	
WP	ACF67367_41	4100001	4210000	
WP	ACF67367_42	4200001	4310000	
WP	ACF67367_43	4300001	4410000	
WP	ACF67367_44	4400001	4510000	
WP	ACF67367_45	4500001	4610000	
WP	ACF67367_46	4600001	4710000	
WP	ACF67367_47	4700001	4810000	
WP	ACF67367_48	4800001	4910000	
WP	ACF67367_49	4900001	5010000	
WP	ACF67367_50	5000001	5110000	
WP	ACF67367_51	5100001	5210000	
WP	ACF67367_52	5200001	5310000	
WP	ACF67367_53	5300001	5410000	
WP	ACF67367_54	5400001	5510000	
WP	ACF67367_55	5500001	5610000	
WP	ACF67367_56	5600001	5648894	
Alignment Scores:				
Pred. No.:	7e-24	Length:	110000	
Score:	271.50	Matches:	54	
Percent Similarity:	72.0%	Conservative:	23	
Best Local Similarity:	50.5%	Mismatches:	25	
Query Match:	49.2%	Indels:	5	
DB:	10	Gaps:	2	
US-10-073-293A-6 (1-111) x ACF67367_20 (1-110000)				
Qy	4	GluValLeuLeuLeuGlyLeuValGlyValAlaAlaAsnTyrCysPheArgTyrLeuPro	23	
Db	22899	AAGATTTCGTGGATTGGACTATTGTTGGTTAGTAACTTTTCATTTCGCTATCTGCCA	22958	
Qy	24	LeuArgLeuArgValGlyAsnAlaArgProThr-----LysArgGlyAlaValGly	40	
Db	22959	CTACGATTT-----GGGAAGACGCCCAATCTGCCGAGAAAGCTGGAAACACAGC	23012	
Qy	41	IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAla	60	
Db	23013	ATTATCCCTTGACAGTATTGGTATTGCATTCATTTCTTTTACTCATCGTATCAGGTGTA	23072	
Qy	61	ProGluValMetHisAspThrArgPheValProThrLeuValGlyPheAlaValLeu	80	

Db	23073	CCTGATGTCGATGAGAGAAAGTCAAAAGTACTTCTTACCCTCATAGGTTGTCTGACCATC	23132	
Qy	81	GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla	100	
Db	23133	TGTTTAGTCTTTTCAAAAACAAAGCAAAATATATACTCGCAACACTATTTGGCGCACTGCTT	23192	
Qy	101	TyGlyLeuAlaTrpLysVal	107	
Db	23193	TTTGACTAAACATTCAAAATA	23213	
RESULT 9				
ACF65382	ID	ACF65382 standard; DNA; 243072 BP.		
XX	AC	ACF65382;		
XX	XX	20-NOV-2003 (first entry)		
DE	XX	Photorhabdus luminescens nucleotide sequence #35.		
XX	XX	Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;		
KW	KW	detection; food; gene expression; plant; animal; microorganism; toxin;		
KW	KW	antibiotic; biopesticide; virulence factor; disease model; plague;		
XX	XX	whooping cough; gene; ds.		
OS	XX	Photorhabdus luminescens.		
XX	XX	WO200294867-A2.		
FN	XX	28-NOV-2002.		
PD	XX			
XX	XX	07-FEB-2002; 2002WO-IB003040.		
PF	XX	07-FEB-2001; 2001FR-00001659.		
XX	XX	(INSP) INST PASTEUR.		
PA	PA	(CNRS) CNRS CENT NAT RECH SCI.		
XX	XX	Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;		
PI	PI	Buchrieser C;		
XX	XX	WPI; 2003-148459/14.		
DR	XX	Genomic sequence of Photorhabdus luminescens and encoded polypeptides,		
XX	XX	useful e.g. as therapeutic antimicrobials and agricultural pesticides.		
PT	XX	Claim 1; SEQ ID NO 35; 1205pp; French.		
XX	XX	The invention relates to the isolation of genes and their encoded		
CC	CC	proteins from Photorhabdus luminescens. The isolated sequences are		
CC	CC	sources of probes and primers for detecting the genome of P. luminescens		
CC	CC	and related species; to study polymorphisms; for gene analysis and for		
CC	CC	detection/amplification of the genes. Antibodies (Ab) raised against the		
CC	CC	polypeptides encoded by the genes are used for detection/identification		
CC	CC	of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that		
CC	CC	carry a gene-containing vector are used to select compounds that		
CC	CC	modulate, regulate, induce or inhibit expression of the genes in plants,		
CC	CC	animals or microorganisms other than P. luminescens and are able to alter		
CC	CC	response or sensitivity to toxins and antibiotics produced by P.		
CC	CC	luminescens. Cells transformed to express the genes are useful for		
CC	CC	recombinant production of the proteins, particularly toxins and		
CC	CC	antibacterials useful as insecticides, bactericides and fungicides. The		
CC	CC	genes, proteins, vectors containing the genes and Ab are also useful		
CC	CC	therapeutically to treat microbial infection by bacteria or fungi that		
CC	CC	are sensitive to P. luminescens-encoded toxins or antibiotics) and as		
CC	CC	biopesticides. Other uses of the genes and the proteins are as virulence		
CC	CC	factors and for identifying targets of human diseases for which P.		
CC	CC	luminescens is a model (particularly plague and whooping cough). This		
XX	XX	sequence represents one of the isolated P. luminescens genes		
XX	XX	Sequence 243072 BP; 70273 A; 57885 C; 49427 G; 65487 T; 0 U; 0 Other;		

PT Escherichia coli K12, useful for detecting gene expression and expression
PT patterns.
XX
PS Claim 3; Page 1834; 2004pp; German.

XX This invention describes a novel biochip comprising probe spots, each
CC containing many identical probes. The probes are nucleotide sequences of
CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at
CC least one includes a segment of at least 20 bases identical with, or
CC complementary to, a segment of an open reading frame (orf) of Escherichia
CC coli K12. The biochip is used for specific detection of gene expression
CC in K12 and for determining the gene expression pattern, e.g. for
CC diagnostic determination of which E. coli strains are present in the gut,
CC and to determine the effects of e.g. growth media on gene expression. The
CC biochip provides as comprehensive as possible detection of the K12
CC genome, with simultaneous analysis of many different genes with a single
CC device, and comparison of gene expression between K12 and its mutants or
CC other E. coli strains in a single experiment. Apart from qualitative and
CC quantitative information about gene expression, it also allows
CC measurements of population densities for the various strains. The use of
CC synthetic oligonucleotides for preparation of probes allows free
CC variation in probe length and ensures high purity (and thus selectivity,
CC reactivity and reproducibility); also synthetic probes are generally
CC shorter than probes prepared by polymerase chain reaction. ACD68731 to
CC ACD81540 represent oligonucleotide probes used with the biochip described
CC in the invention

SQ Sequence 100 BP; 15 A; 31 C; 29 G; 25 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.69e-13 Length: 100
Score: 164.00 Matches: 33
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 29.7% Indels: 0
DB: 8 Gaps: 0

US-10-073-293A-6 (1-111) x ACD80498 (1-100)

Qy 50 SerTleCysAlaLeuLeuValValSerThrAlaProGluValMetHisAspThrArgArg 69
Db 1 TCGATATGCGCTCTGCTGTTGTCTCTACCGCAGAGAGTGATGCACGATACACGCGGT 60
Qy 70 PheValProThrLeuValGlyPheAlaValLeuGlyAla 82
Db 61 TTCGTGCCCCACGCTGCTGCGCTTCGCGTACTGGGTGCC 99

RESULT 12
ACD80497

ID ACD80497 standard; DNA; 100 BP.

XX ACD80497;

XX 19-SEP-2003 (first entry)

XX E. coli K12 MG1655 biochip probe SEQ ID 11773.

XX Biochip; gene expression; gut; diagnostic; detection; probe; ss.

XX Escherichia coli.

XX EPI260592-A1.

XX 27-NOV-2002.

XX 17-MAY-2001; 2001EP-00112179.

XX 17-MAY-2001; 2001EP-00112179.

XX (MWGB-) MWG-BIOTECH AG.

XX Donner H, Drescher B, Huber A, Weber J;

DR WPI; 2003-241155/24.

XX Biochip containing probes complementary with open reading frames in
PT Escherichia coli K12, useful for detecting gene expression and expression
PT patterns.

PS Claim 3; Page 1834; 2004pp; German.

XX This invention describes a novel biochip comprising probe spots, each
CC containing many identical probes. The probes are nucleotide sequences of
CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at
CC least one includes a segment of at least 20 bases identical with, or
CC complementary to, a segment of an open reading frame (orf) of Escherichia
CC coli K12. The biochip is used for specific detection of gene expression
CC in K12 and for determining the gene expression pattern, e.g. for
CC diagnostic determination of which E. coli strains are present in the gut,
CC and to determine the effects of e.g. growth media on gene expression. The
CC biochip provides as comprehensive as possible detection of the K12
CC genome, with simultaneous analysis of many different genes with a single
CC device, and comparison of gene expression between K12 and its mutants or
CC other E. coli strains in a single experiment. Apart from qualitative and
CC quantitative information about gene expression, it also allows
CC measurements of population densities for the various strains. The use of
CC synthetic oligonucleotides for preparation of probes allows free
CC variation in probe length and ensures high purity (and thus selectivity,
CC reactivity and reproducibility); also synthetic probes are generally
CC shorter than probes prepared by polymerase chain reaction. ACD68731 to
CC ACD81540 represent oligonucleotide probes used with the biochip described
CC in the invention

SQ Sequence 100 BP; 18 A; 31 C; 24 G; 27 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.93e-13 Length: 100
Score: 160.00 Matches: 33
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 29.0% Indels: 0
DB: 8 Gaps: 0

US-10-073-293A-6 (1-111) x ACD80497 (1-100)

Qy 69 ArgPheValProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrLysThrArg 88
Db 2 CGTTTCGTGCCCCACGCTGCTGCGCTTCGCGTACTGGGTGCCACTTCTATATAACACGC 61
Qy 89 SerIleIleIleProThrLeuLeuSerAlaLeuAlaTyr 101
Db 62 AGCATTATCATCCCAACACTGCTTAGTGGCGCTGGCCTAT 100

RESULT 13

ACD80499

ID ACD80499 standard; DNA; 100 BP.

XX ACD80499;

XX 19-SEP-2003 (first entry)

XX E. coli K12 MG1655 biochip probe SEQ ID 11775.

XX Biochip; gene expression; gut; diagnostic; detection; probe; ss.

XX Escherichia coli.

XX EPI260592-A1.

XX 27-NOV-2002.

XX 17-MAY-2001; 2001EP-00112179.

XX 17-MAY-2001; 2001EP-00112179.

XX (MWGB-) MWG-BIOTECH AG.

XX Donner H, Drescher B, Huber A, Weber J;
XX WPI; 2003-241155/24.
XX Biochip containing probes complementary with open reading frames in
XX Escherichia coli K12, useful for detecting gene expression and expression
XX patterns.
XX Claim 3; Page 1834; 2004pp; German.
XX This invention describes a novel biochip comprising probe spots, each
XX containing many identical probes. The probes are nucleotide sequences of
XX 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at
XX least one includes a segment of at least 20 bases identical with, or
XX complementary to, a segment of an open reading frame (orf) of Escherichia
XX coli K12. The biochip is used for specific detection of gene expression
XX in K12 and for determining the gene expression pattern, e.g. for
XX diagnostic determination of which E. coli strains are present in the gut,
XX and to determine the effects of e.g. growth media on gene expression. The
XX biochip provides as comprehensive as possible detection of the K12
XX genome, with simultaneous analysis of many different genes with a single
XX device, and comparison of gene expression between K12 and its mutants or
XX other E. coli strains in a single experiment. Apart from qualitative and
XX quantitative information about gene expression, it also allows
XX measurements of population densities for the various strains. The use of
XX synthetic oligonucleotides for preparation of probes allows free
XX variation in probe length and ensures high purity (and thus selectivity,
XX reactivity and reproducibility); also synthetic probes are generally
XX shorter than probes prepared by polymerase chain reaction. ACD8731 to
XX ACD81540 represent oligonucleotide probes used with the biochip described
XX in the invention

SQ Sequence 100 BP; 18 A; 29 C; 26 G; 27 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.85e-12 Length: 100
Score: 155.00 Matches: 32
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 28.1% Indels: 0
DB: 8 Gaps: 0

US-10-073-293A-6 (1-111) x ACD80499 (1-100)

Qy 38 AlaValGlyIleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuValVal 57
Db 3 GCGTAGGTATTTTGGTCGACACCATTTGGCATCGCTCGATATGCGCTCTGCTGTTGTC 62
Qy 58 SerThrAlaProGluValMetHisAspThrArgArg 69
Db 63 TCTACCGCACCAGAGTGATGCACGATACACGCCGT 98

RESULT 14
ACF79478
ID ACF79478 standard; DNA; 172 BP.
XX ACF79478;
XX 18-DEC-2003 (first entry)
XX Escherichia coli multidrug transporter emrR gene promoter.
XX Promoter; emrR gene; multidrug transporter; ds.
XX Escherichia coli.
XX WO2003066853-A1.
XX 14-AUG-2003.
XX 07-FEB-2003; 2003WO-US003902.
XX

PR 08-FEB-2002; 2002US-0355083P.
PR 13-JUN-2002; 2002GB-00013616.
XX (AVET) AVENTIS PHARM INC.
XX Macneil I, Tiong CUY, Brown K;
XX WPI; 2003-731503/69.
XX Identifying a DNA molecule that encodes, or a protein involved in the
XX production of, a natural product having bioactivity by incubating the
XX transformed bacterial cell under conditions that permit the production of
XX the natural product.
XX Disclosure; Page 20; Opp; English.

CC The present sequence is the promoter of the Escherichia coli multidrug
CC transporter protein gene emrR. The promoter can be used in the method of
CC the invention for identifying a DNA molecule that encodes a natural
CC product having bioactivity, or a protein involved in the production of
CC the natural product. The method involves: (a) providing a transformed a
CC bacterial cell comprising (i) a gene fusion encoding a reporter protein
CC associated with a homeostatic promoter which is activated by the natural
CC product, and (ii) an exogenous DNA that encodes the natural product or
CC the protein involved in production of the natural product; (b) incubating
CC the transformed bacterial cell under conditions that permit production of
CC the natural product; and (c) detecting the presence of the reporter
CC protein within the transformed bacterial cell

SQ Sequence 172 BP; 49 A; 37 C; 32 G; 54 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.00037 Length: 172
Score: 98.00 Matches: 20
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 17.8% Indels: 0
DB: 10 Gaps: 0

US-10-073-293A-6 (1-111) x ACF79478 (1-172)

Qy 92 IleProThrLeuLeuSerAlaLeuAlaTyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 2 ATCCCAACACTGCTTAGTGGCTGCTATGGCTCGCTCGCTGGAAGTGATGGCGATTATA 61

RESULT 15
ACA39729
ID ACA39729 standard; DNA; 1593 BP.
XX ACA39729;
XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #21386.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX Mycobacterium leprae.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU35859.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 27599; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1593 BP; 239 A; 469 C; 567 G; 318 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.61 Length: 1593
 Score: 78.00 Matches: 31
 Percent Similarity: 40.5% Conservative: 14
 Best Local Similarity: 27.9% Mismatches: 32
 Query Match: 14.1% Indels: 34
 DB: 8 Gaps: 5

US-10-073-293A-6 (1-111) x ACA39729 (1-1593)

QY 4 GluValLeu-----LeuLeuGlyLeuValGlyValAlaAlaAsnTyrCysPheArg 20
 Db 166 GAGGTGCTGTCGGGTGATCTGTTCGTCATTCGTCGTCATCCACACGGCGCTCGG 225
 QY 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
 Db 226 TATCCCGCGTTCGCTC-----GAACGAGGTGCTGTCGCG 261
 QY 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuValValSerThrAla 60
 Db 262 GTGCTTACCGACGTTCGAGGGTCCCGAGCTGACAGCC-----CAGGCCAGCAGCGCTG 315
 QY 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
 Db 316 CCGATACCTCATACATCCGGAACCCCGC-----AGCGTGCTC 351

Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleIleProThrLeuLeuSerAlaLeuAla 100
 Db 352 GGCAGT-----CTTGCCTCCGCGCA 372
 Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
 Db 373 TATGACATCCCTCCAAACCGGATGACGGTTGTT 405

Search completed: February 16, 2006, 18:26:27
 Job time : 350.935 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:12:59 ; Search time 2260.53 Seconds
(without alignments)
2297.407 Million cell updates/sec

Title: US-10-073-293A-6

Perfect score: 552
Sequence: 1 MSYEVLLGLLVGVNYCFR.....IPTLLSALAYGLAWKVMAIL 111

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/absas/ABSSWEB.spool/US10073293/runat.15022006.120736.1008/app.query.fasta_1
-DB=EST -QPMW=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs04
-USER=US10073293 @CGN 1.1 6731 @runat.15022006.120736.1008 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gses1.*
10: gb_gses2.*
11: gb_gses3.*

754 bp mRNA linear EST 08-MAY-2001
602694488F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4826757 5',
mRNA sequence.

Accession: BG721873.1 GI:14001060
Version: BG721873.1
Keywords: EST.
Source: Homo sapiens (human)
Organism: Homo sapiens

Reference: NIH-MGC http://mgc.nci.nih.gov/
Authors: National Institutes of Health, Mammalian Gene Collection (MGC)
Title: Unpublished (1999)
Journal: Contact: Robert Strausberg, Ph.D.
Comment: Email: cgapbs@mail.nih.gov

1 (bases 1 to 754)
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov
Plate: L1M10741 row: p column: 22
High quality sequence stop: 691.

Location/Qualifiers
1. .754
source

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
8	85	15.4	755	7	CV111390	CV111390	AGENCOURT					
9	83.5	15.1	547	8	DN993155	DN993155	TC102327					
10	83	15.0	584	3	BI972603	BI972603	sai10c10.					
11	82.5	14.9	431	2	BB844978	BB844978	BB844978					
12	82.5	14.9	717	1	AJ448384	AJ448384	AJ448384					
13	82.5	14.9	728	8	DR426972	DR426972	new37a01.					
14	82.5	14.9	730	1	AJ446899	AJ446899	AJ446899					
15	82.5	14.9	788	1	AJ446332	AJ446332	AJ446332					
16	82.5	14.9	811	1	AJ452083	AJ452083	AJ452083					
17	82.5	14.9	827	1	AJ455413	AJ455413	AJ455413					
18	82.5	14.9	958	5	BU421186	BU421186	603954658					
19	82	14.9	534	1	AL681632	AL681632	AL681632					
20	82	14.9	555	1	AL849287	AL849287	AL849287					
21	82	14.9	661	1	AL856042	AL856042	AL856042					
22	82	14.9	681	9	BZ088402	BZ088402	11a99a03.					
23	82	14.9	706	7	CN102523	CN102523	EC2CAAIDE					
24	82	14.9	813	1	AL965519	AL965519	AL965519					
25	82	14.9	836	7	CN095349	CN095349	EC2BAALDE					
26	82	14.9	843	8	CX442865	CX442865	JGI_XZ684					
27	82	14.9	847	9	BH465150	BH465150	BOG0230TR					
28	82	14.9	1701	10	AY408608	AY408608	Mus muscu					
29	82	14.9	2005	2	BF034156	BF034156	601456525					
30	82	14.9	2160	4	AK016870	AK016870	Mus muscu					
31	81.5	14.8	432	6	CB385955	CB385955	OSTF036G1					
32	81	14.7	767	7	CK599023	CK599023	AGENCOURT					
33	80.5	14.6	594	3	BJ675006	BJ675006	BJ675006					
34	80.5	14.6	651	2	BF337569	BF337569	602035320					
35	80.5	14.6	818	2	BG720995	BG720995	602692691					
36	80.5	14.6	947	5	BQ652127	BQ652127	AGENCOURT					
37	79.5	14.4	499	11	TA369G01P	TA369G01P	AL463356 T. brucei					
38	79.5	14.4	667	6	CD475939	CD475939	na003-12m					
39	79.5	14.4	770	2	BF121899	BF121899	601758733					
40	79	14.3	864	7	CO733249	CO733249	SLT02C04					
41	78.5	14.2	590	8	DN956994	DN956994	Fh.mxo.25					
42	78.5	14.2	591	7	CN303248	CN303248	170006000					
43	78.5	14.2	738	7	CN303263	CN303263	170005999					
44	78.5	14.2	918	10	CG004411	CG004411	Z00A034TH					
45	78.5	14.2	944	7	CO583414	CO583414	ILLUMINGEN					

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91.5	16.6	754	2	BG721873 602694488
2	86.5	15.7	511	2	BG881410 sae1g11.
3	86.5	15.7	776	8	DR425028 new13g11.
4	86.5	15.7	897	5	BU375062 603585638
5	86.5	15.7	916	5	BU375599 603812270
6	86.5	15.7	958	5	BU274235 603532828
7	85.5	15.5	651	7	CN847401 PG07029C0

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:482675"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer_
5'-TTTTTTTTTTTTTN-3', size-selected for average
insert size 2.2 kb and normalized to R0T 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      0.542      Length:      754
Score:          91.50      Matches:    33
Percent Similarity: 50.0%      Conservative: 24
Best Local Similarity: 28.9%      Mismatches: 36
Query Match:      16.6%      Indels:    21
DB:               2         Gaps:      8

```

US-10-073-293A-6 (1-111) x BG721873 (1-754)

```

Qy      2  SerTyrGluValLeuLeuGly----LeuLeuValGlyVal----AlaAsnTyrCysPhe 19
|||
|||
|||
440  TCAGTTGAATGCTGATCTGGCGCGCTGTATTGGCCCTCTTCTGGGACTCTGCACA 499
Db

Qy      20 ArgTyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLys---ArgGlyAla 38
|||
|||
|||
500  GGTTTTGGCCC-----ATGTACATTGGAGAGATCTGCCTACTGCGCTGAGGGGTGCC 553
Db

Qy      39 ValGlyLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuValValSer 58
|||
|||
|||
554  TTGGC---ATCTCAACACAGCTGGGCATA-----GTATTGGAATTCTGGTGGCCAG 604
Db

Qy      59 Thr---AlaProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPhe 77
|||
|||
|||
605  ATCTTTGCTGTGAACCTATCTCTGGGCTCGAAGAGCTATGCGCGGTCTATTAGGCTTT 664
Db

Qy      78 AlaValLeuGlyAlaSerPheTyrLysThrArgSerIleIlelleProThrLeuLeuSer 97
|||
|||
|||
665  ACCATCTCTT-----CCAGCTATCTCTGCAA 688
Db

Qy      98 AlaLeuAlaTyrGlyLeuAlaTrpLysValMetAlaIleIle 111
|||
|||
|||
689  AGTCAGCCCTCCATGTTGGCCCTGAAGTCCAGATTGTC 730
Db

```

RESULT 2

```

BG881410
LOCUS      BG881410          511 bp      mRNA      linear      EST 24-JUL-2004
DEFINITION  sa881g11.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl065-2613 5' similar to TR:Q9XIR3 Q9XIR3 F13011.12 PROTEIN.
; mRNA sequence.

```

ACCESSION

VERSION BG881410

KEYWORDS

EST.

SOURCE

Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE

1 (bases 1 to 511)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,

TITLE
JOURNAL
COMMENT

McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Other_ESTs: BU548976 corresponding to Gm-rl088-5925 (3')
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert length: 1836 Std Error: 0.00
High quality sequence stop: 421.

FEATURES

Location/Qualifiers

1..511
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-2613"
/tissue_type="germinating shoots"
/lab_host="DH10B"
/clone_lib="Gm-cl065"
/note="vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
germinating shoots of the cultivar Williams. The seeds
were allowed to germinate for 24 hours prior to being
cold stressed for 2 days at 4C. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."

ORIGIN

```

Alignment Scores:
Pred. No.:      1.44      Length:      511
Score:          86.50      Matches:    34
Percent Similarity: 47.5%      Conservative: 24
Best Local Similarity: 27.9%      Mismatches: 40
Query Match:      15.7%      Indels:    25
DB:               2         Gaps:      4

```

US-10-073-293A-6 (1-111) x BG881410 (1-511)

```

Qy      2  SerTyrGluValLeuLeuGlyLeuVal---GlyValAlaAsnTyrCysPheArg 20
|||
|||
|||
105  TCTCTGGGTGTCTATTGCTACCTCTTTGCTGAAGTCACGGATTACGGCTCTTTGG 164
Db

Qy      21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
|||
|||
|||
165  GAATGTCGGTGCATTGGAACTTTATGTGGCAACCTATGTGGCGGAGATATAGGA 224
Db

Qy      41 IleLeuLeuAspThrIleGlyIle----- 48
|||
|||
|||
225  AGTTTGGAAACTACATTGGTGGATGTGACTGTGACATTGGTTGTCCTCCGATAT 284
Db

Qy      49 -----AlaSerIleCysAlaLeuValValSerThrAlaProGluVal 63
|||
|||
|||
285  CTCTGGTGTGGATATTCACTGATAAA-ATACTGTGTGTTTACTCAAGACCTGAGATT 343
Db

Qy      64 MethHisAspThrArgArgPhe-----ValProThrLeuValGlyPheAlaVal 79
|||
|||
|||
344  TCTCATGAGCTCGTGAGTACTGCATATACCTCATTCAGCTCTATTGTTGGTCATGCTGTT 403
Db

```



```

Db      615 CTTGTTTGGTGGGA 629
|||||
RESULT 8
CV111390
LOCUS
DEFINITION CV111390 755 bp mRNA linear EST 30-AUG-2004
IMAGE:7457639 5', mRNA sequence.
ACCESSION CV111390
VERSION
KEYWORDS
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 755)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILLNL at:
http://image.llnl.gov
Plate: LLN15723 row: d column: 21
High quality sequence stop: 703.
Location/Qualifiers
1. 755
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7457639"
/tissue_type="testis, pooled"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 238"
/notes="Organ: testis; Vector: pExpress-1; Site 1: EcoRV;
Site 2: NotI; RNA obtained from testis tissue of 8 wk old
animal. Tissues were snap-frozen and kept at -80C before
RNA extraction and purification (Tri-reagent method). cDNA
was primed using oligo-dT primer:
5'-pGACTAGTCTAGATCGAGCGCGCC(T)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb
resulted in an average insert size of 1.9 kb. This primary
library is normalized (non-normalized primary library is
NIH MGC 237) and was constructed by Express Genomics
(Frederick, MD)"
ORIGIN
Alignment Scores: 3.74 Length: 755
Pred. No.: 85.00 Matches: 31
Score: 43.9% Conservative: 12
Best Local Similarity: 31.6% Mismatches: 39
Query Match: 15.4% Indels: 16
DB: 7 Gaps: 5

US-10-073-293A-6 (1-1111) x CV111390 (1-755)

Qy      18 CysPheArgTyrLeuPro---LeuArgLeuArgValGlyAsnAlaArgProThrIysArg 36
|||||
Db      231 TGTGGAGGCTCAGGCCACGACGATGACACGGGTTCCGGCGCTCGGCCCTCGGAAGAGA 290
|||||
Qy      37 GlyAlaVal-----GlyIleLeuLeuAspThrIleGlyIleAlaSerIle 51
|||||
Db      291 GCTGCCATCTTAACATACGAGGCGCTGTATTTCCAGGCTTTCGGAGGACGCGCGC 350
|||||

```

```

Qy      52 CysAlaLeuLeuValValSerThrAlaProGluValMetHisAspThrArgAlaGpPheVal 71
|||||
Db      351 TGC-----CTCACAAGAGTGTGCCCCCATGCTAGCCCATGGCTACCGCGCTTCTCTA 404
|||||
Qy      72 ProThrLeuValGlyPheAlaValLeu-----GlyAlaSerPheTyrLysThrArg 88
|||||
Db      405 CCTAGGAAGGACACCGTGTTCATCTTCGACTATGTGGGGCCCTCTTCTTCTCAAAAT 464
|||||
Qy      89 SerIleIleLeuProThrLeuLeuSerAlaLeuAlaTyrGlyLeuAlaTrpLys 106
|||||
Db      465 GCTCTGCTGCTCTCCACCTTGGCCAG-----ATCCAGTGAAG 503
|||||

RESULT 9
DN993155
LOCUS
DEFINITION DN993155 547 bp mRNA linear EST 17-MAY-2005
TCl02327 Human adult whole brain, large insert, pCMV expression
library Homo sapiens cDNA clone TCl02327 5' similar to Homo sapiens
cDNA FLJ38307 fis, clone FCBBF3018949, mRNA sequence.
ACCESSION DN993155
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,
Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,
Zhang,X., Jay,G. and He,W.
TITLE High-throughput cloning of full-length human cDNAs directly from
cDNA libraries optimized for large and rare transcripts
JOURNAL Unpublished (2005)
COMMENT Contact: Kovacs, KF
High Throughput cDNA Cloning
OriGene Technologies, Inc. ( www.oriGene.com )
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cDNA@origene.com
This EST submission is part of an on-going human full-length
cloning project at OriGene Technologies, Inc.
Please contact OriGene for access.
OriGene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.oriGene.com
Seq primer: pCMV6 5prime forward vector primer, OriGene
Technologies Inc.
Location/Qualifiers
1. 547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCl02327"
/tissue_type="Whole brain"
/clone_lib="Human adult whole brain, large insert, pCMV
expression library"
/notes="Organ: Brain; Vector: pCMV6-XL5; Site 1: EcoRI;
Site 2: XhoI/SalI compatible end ligatio; Oligo-dT primed
reverse transcription optimized for large and GC rich mRNA
transcripts, cDNA size selection, optimized ligation for
large inserts into mammalian expression vector, random
clones selected for end sequence verification of
full-length genes"
ORIGIN
Alignment Scores: 3.83 Length: 547
Pred. No.: 83.50 Matches: 34
Score: 40.2% Conservative: 15
Percent Similarity:

```


Best Local Similarity: 27.9% Mismatches: 44
Query Match: 15.1% Indels: 29
DB: 8 Gaps: 6

US-10-073-293A-6 (1-111) x DN993155 (1-547)

Qy 7 LeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArgTyrLeuProLeuArgLeu 26
|||||
Db 119 CTTTGGCTTCTCACCTGGGGCGGTGTTCTGCACACCTCCACCTCTCTCTC 178
|||||
Qy 27 ArgValGlyAan-----
|||||
Db 179 CGTGGACGATAGACGACATCCAGACACGACAGCTCCCGGGCGCTTCTATCTC 238
|||||
Qy 34 ThrLysArgGlyAlaValGlyIleLeuLeuAspThrIleGlyIleAlaSer----- 50
|||||
Db 239 CGTCTCCAGGACGACACA-----AGCCATTGTGGAACTCTCCACCGGTGTACA 286
|||||
Qy 51 -----lleCysAlaLeu---LeuValValSerThrAlaProGluValMet----- 64
|||||
Db 287 GAACGGTGCCTCTCGTCCCTGCCACTCAGGACCTCTCAAGTCCCGGATGTGATGGCTCCT 346
|||||
Qy 65 ---HisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeuGlyAlaSer 83
|||||
Db 347 CAGCATGATCAGGAGAAATCCATGAT-----CTTGTATTCTCTCTTGGGAAGTCC 400
|||||
Qy 84 PheTyrLysThrArgSerIleIleProThrLeuLeuSerAlaLeuAlaTyrGlyLeu 103
|||||
Db 401 TTCTCCATGCTTAACCAAGATCTATATGGCTATAGCACCAGCTCTTGGCTCTTGGCTTG 460
|||||
Qy 104 AlaTrp 105
|||||
Db 461 GCATGG 466
|||||

RESULT 10
BI972603/c
LOCUS
DEFINITION
sa180c10.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl065-7052 5' similar to TR:Q9XI52 Q9XI52 F9L1.11 PROTEIN.
; mRNA sequence.

ACCESSION BI972603
VERSION BI972603.1 GI:16347008
KEYWORDS
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1. (bases 1 to 584)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, S., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R., and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. Possible
reversed clone: similarity on wrong strand this clone is available
through: Bogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA
(phone: 800 423 4163; email: info@biogeneticservices.com)
High quality sequence stop: 391.
Location/Qualifiers
1. 584

FEATURES
Source

/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-7052"
/tissue_type="germinating shoots"
/lab_host="DH10B"
/clone_lib="Gm-cl065"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
germinating shoots of the cultivar Williams. The seeds
were allowed to germinate for 24 hours prior to being
cold stressed for 2 days at 4C. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."

ORIGIN

Alignment Scores:
Pred. No.: 4.84 Length: 584
Score: 83.00 Matches: 29
Percent Similarity: 45.7% Conservative: 24
Best Local Similarity: 25.0% Mismatches: 35
Query Match: 15.0% Indels: 29
DB: 3 Gaps: 4

US-10-073-293A-6 (1-111) x BI972603 (1-584)

Qy 23 ProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGlyLeu 42
|||||
Db 532 CCGGTGCATGGAACCTTATGTGGCAACCTATGTGGGAAGAGTATAGGAAGTTG 473
|||||
Qy 43 LeuAspThrIleGlyIle-----
|||||
Db 472 GAAATTACACTTGGTGGCGATTGTGACTGTGACATTGGTTGCTCCGATATCTCTGG 413
|||||
Qy 49 -----AlaSerIleCysAlaLeuLeuValSerThrAlaProGluValMetHis 65
|||||
Db 412 TGTGGATATTCACTGATAAA-ATACTGTCTGTTAGTCAAGACCCTGAGATTCTCAT 354
|||||
Qy 66 AspThrArgArgPhe-----ValProThrLeuValGlyPheAlaValLeuGly 81
|||||
Db 353 GCAGCTCGTGAGTACTGCATATACCTCATTCAGCTCTATTGGTCATGCTGTTCTCAA 294
|||||
Qy 82 Ala-----SerPheTyrLysThrArgSerIleIleProThrLeuLeuSerAla--- 98
|||||
Db 293 GCTCTACTCGTACTTCCAGATCAGATGATGATCTTCCCATGGTTTCAGCTCAATC 234
|||||
Qy 99 -----LeuAlaTyrGlyLeuAlaTyrLysValMetAlaIle 110
|||||
Db 233 ACCGCACTGGTTTGCATGTTCTCTATTGTTGGAGTCTCTGATTGGTT 186
|||||

RESULT 11

BB844978
LOCUS
DEFINITION
BB844978 RIKEN full-length enriched, adult male kidney Mus musculus
cDNA clone F520002L02 5', mRNA sequence.

ACCESSION BB844978.1 GI:17083353
VERSION BB844978 EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Muridae; Mus.
1 (bases 1 to 431)
REFERENCE
AUTHORS
Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,


```

US-10-073-293A-6 (1-111) x AJ446899 (1-730)

Qy 5 ValLeuLeuLeuGlyLeuLeuValGlyValAlaAsnTyrCysPheArgTyrLeuProLeu 24
Db 688 GTCTTGCTGGTGTGCTGCTGCTTTGTTGGTGTGTT-----ATCTTGTGGGTTGCTGTT 638
Qy 25 ArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGlyIleLeuLeuAsp 44
Db 637 GCTTTGCTGGTGTGCTGCTGCTGCTGGCACCCCACT---GTTGCTCTTTCGGCCATTGTTG 581
Qy 45 ThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAlaProGluValMet 64
Db 580 GCCATCGGTGTGTCATCTTCTCCTGCCATTGTTGCCATCAGTGTGCCATTGTCGCGT 521
Qy 65 HisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeuGlyAlaSerPhe 84
Db 520 CTGACTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
Qy 85 TyrIlysthrArgSerIleIleProThrLeuLeuSerAlaLeuAlaTyr----- 101
Db 460 GTGGGTGTTTGGCCGTGATCATCACCACCTGGGTGTTGCTTGTGGCTGTCACGTGTTCTG 401
Qy 102 -----GlyLeuAlaTyrLysValMetAla 109
Db 400 GGTGCTGGCTGGCTAGCAGTGTGCTGCC 371

RESULT 15
AJ446332/c
LOCUS AJ446332 788 bp mRNA linear EST 19-APR-2002
DEFINITION AJ446332 riken1 Gallus gallus cdna clone 13f14r1, mRNA sequence.
ACCESSION AJ446332
VERSION AJ446332.1 GI:20213553
KEYWORDS EST.
SOURCE
ORGANISM Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Buerstedde,J.M.
Gallus gallus bursal lymphocyte EST
Unpublished (2002)
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

FEATURES
Source
1..788
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="13f14r1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/clone_lib="riken1"
/note="CB inbred strain"

ORIGIN
Alignment Scores:
Pred. No.: 8.31 Length: 788
Score: 82.50 Matches: 37
Percent Similarity: 44.5% Conservative: 12
Best Local Similarity: 33.6% Mismatches: 52
Query Match: 14.9% Indels: 9
DB: 1 Gaps: 3

US-10-073-293A-6 (1-111) x AJ446332 (1-788)

Qy 5 ValLeuLeuLeuGlyLeuLeuValGlyValAlaAsnTyrCysPheArgTyrLeuProLeu 24
Db 695 GTCTTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645

```

```

Qy 25 ArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGlyIleLeuLeuAsp 44
Db 644 GCTTTGCTGGTGTGCTGCTGCTGCTGGCACCCCACT---GTTGCTCTTTCGGCCATTGTTG 588
Qy 45 ThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAlaProGluValMet 64
Db 587 GCCATCGGTGTGTCATCTTCTCCTGGCCATTGTTGCCATCAGTGTGCCATTGTCGCGT 528
Qy 65 HisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeuGlyAlaSerPhe 84
Db 527 CTGACTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468
Qy 85 TyrIlysthrArgSerIleIleProThrLeuLeuSerAlaLeuAlaTyr----- 101
Db 467 GTGGGTGTTTGGCCGTGATCATCACCACCTGGGTGTTGCTTGTGGCTGTCACGTGTTCTG 408
Qy 102 -----GlyLeuAlaTyrLysValMetAla 109
Db 407 GGTGCTGGCTGGCTAGCAGTGTGCTGCC 378

```

Search completed: February 16, 2006, 21:44:14
Job time : 2265.53 secs

Qy 21 TyrLeuProLeuArgLeuArgValGlyAenAlaArgProThrLysArgGlyAlaValGly 40
Db 97 TATCTTCTCTGCGTTTGGCGCGCCGACGCTGCTCTGCGCCGCGCGCGTCAAGC 156
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAla 60
Db 157 GTGCTGCTGGATATATTGGATCGCTCGATTTGGCGGCTGCTGGTGTCTCCAGCGTG 216
Qy 61 ProGluValMetHisAspThrArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 217 CCGGAGATCTTCGCGCATGCCATCGCTGCGCGCGAGCTGACCGGCTTCTCGTGCTG 276
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleProThrLeuSerAlaLeuAla 100
Db 277 GGGCGCGCTCTTGGAAACCCGAGCATCATTTGTCGGACCTGCTGCGCGCTTTGCC 336
Qy 101 TyrGlyLeuAlaTPrLysVal 107
Db 337 TATGCTCTGGCATGGAAATC 357

RESULT 2

US-09-328-352-2738
; Sequence 2738, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09-328-352
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2738
; LENGTH: 369
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii

US-09-328-352-2738

Alignment Scores:
Pred. No.: 3.84e-15 Length: 369
Score: 182.50 Matches: 39
Percent Similarity: 63.5% Conservative: 34
Best Local Similarity: 33.9% Mismatches: 35
Query Match: 33.1% Indels: 7
Gaps: 3

US-10-073-293A-6 (1-111) x US-09-328-352-2738 (1-369)

Qy 1 MetSerTyrGluValLeuLeuLeuGlyLeuLeuValGlyValAlaAAsnTyrCysPheArg 20
Db 28 ATGAACCTAGAAATATTATTTGGTCGGCATTAATGTGGGTATTTGCTAAATTTTGTTCACGC 87
Qy 21 TyrLeuPro-----LeuArgLeuArgValGlyAenAlaArgProThrLysArgGly 37
Db 88 TCGGACCATTTTCGTAAATCAAAACTACAGGCAACCAACAA-----AGAGGTGC 141
Qy 38 AlaVal-----GlyIleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeu 55
Db 142 TCAATCTGGTTTAAATAATCGCTTAGGAAGTATTGGTATATCTGCAATTTAGCGCCATGCTT 201
Qy 56 ValValSerThrAlaProGluValMetHisAspThrArgPheValProThrLeuVal 75
Db 202 GTGGTTGCAACACTTCCACCTTTACTTGAACCCCAATAAGAGTTTAGCAATGCTTATC 261
Qy 76 GlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeu 95
Db 262 GGTATTTAGTTTAGCAGGCTTTATTTAAATTTAAGAAATGTTACCAAGCACTTTA 321
Qy 96 LeuSerAlaLeuAlaTPrLysValMetAlaIle 110
Db 322 ACTGCCGAATCGTTTATGGCCTTATCTATATCTTATTTACCTTTA 366

RESULT 3

US-09-596-002-10
; Sequence 10, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US/09/596,002
; CURRENT APPLICATION NUMBER: 2000-06-16
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 19988
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 10
; PUBLICATION INFORMATION:
US-09-596-002-10

Alignment Scores:

Pred. No.: 86.1 Length: 19988
Score: 77.50 Matches: 32
Percent Similarity: 40.3% Conservative: 20
Best Local Similarity: 24.8% Mismatches: 36
Query Match: 14.0% Indels: 41
Gaps: 6

US-10-073-293A-6 (1-111) x US-09-596-002-10 (1-19988)

Qy 9 GlyLeuLeuValGlyValAlaAAsnTyrCysPheArgTyr----- 21
Db 2843 GGTGCGCTCATCGGCTTAATCAGTGTTCGACCGCTTTTACCATCTTTGCCAACATTGAT 2902
Qy 22 -----LeuProLeuArgLeuArgValGlyAenAlaArgProThrLys 35
Db 2903 TTGATTTTAGATGCTGACTTACCCATGCAATATTGTGGCAATGTTTCATCCATGG--- 2959
Qy 36 ArgGlyAlaValGlyIleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeu 55
Db 2960 -----CTTGGCACTTTGATGAGTTTATCATCTTTTGGC-----ATG 2995
Qy 56 ValValSerThrAlaProGluValMetHis-----Asp 66
Db 2996 ATTTTAAATACCGCCATCAGTTTATATCTGCTGCGCAAGCGTTTTTCAGGTGATGAC 3055
Qy 67 ThrArgArgPheValProThrLeuValGlyPheAlaValLeuGly----- 81
Db 3056 AACGAGCGCTTAAATGGATATTGATTTGGCTTGTGTTTGTGCGCTTTATCTTAAGCTTT 3115
Qy 82 AlaSerPheTyrLysThrArgSerIleIleIleProThr-----LeuLeu 96
Db 3116 GCAGTTTAAATAAACGTGCTCAATCATGTTCCCATCATTTGGCTACATTTGGTATGTTG 3175
Qy 97 SerAlaLeuAlaTyrGlyLeuAlaTPr 105
Db 3176 CTTCTTCTGCGTGTGTTGTTGGCATGG 3202

RESULT 4

US-09-711-164-57
; Sequence 57, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERET

```

; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711.164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-711-164-57

Alignment Scores:
Pred. No.: 0.313 Length: 330
Score: 77.00 Matches: 24
Percent Similarity: 41.9% Conservative: 20
Best Local Similarity: 22.9% Mismatches: 37
Query Match: 13.9% Indels: 24
DB: 3 Gaps: 3

US-10-073-293A-6 (1-111) x US-09-711-164-57 (1-330)

Qy 24 LeuArgLeuValGlyAsnAlaArgProThrLysArgGlyAlaValGlyLeuLeu 43
Db 22 GTGGCTTCAGGATCGCATCTGCAGACCTTTTCACGCGGCATAATGCGCTTCATGCTC 81
Qy 44 -----AspThrIleGlyIleAlaSerIle 51
Db 82 GGGCCTTCGTCTACGAAATTTTCGTAACTTTTCAGATCGTCAATGTCCAGCGCCATCGTTG 141
Qy 52 CysAlaLeuLeuValSerThrAlaProGluValMethHisAspThrArgPheVal 71
Db 142 TGTTCAGCGTTA-----GCAATGGCAGATTCAGAACTTTCTTG 180
Qy 72 ProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrlsThrArgSerIle 91
Db 181 ACCAGTACAGCCGCTTTCTTGTGGTGTAGGTCAAAATATATCCAGAGCTCGCACACTTTC 240
Qy 92 IleProThrLeuSerAlaLeu-----AlaTyrlsGlyLeuAlaTrpLys 106
Db 241 TTACCGGATCAGGTTCAGTACGACACGCGGACCTTCTGACGAGNAGAACGAGCATGGCGA 300
Qy 107 ValMetAlaIleIle 111
Db 301 TGTTCAGCATAGT 315

RESULT 5
US-09-711-164-134/c
; Sequence 134, Application US/09/711.164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Alllyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711.164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(333)
US-09-711-164-134

Alignment Scores:
Pred. No.: 0.317 Length: 333
Score: 77.00 Matches: 24
Percent Similarity: 41.9% Conservative: 20
Best Local Similarity: 22.9% Mismatches: 37
Query Match: 13.9% Indels: 24
DB: 3 Gaps: 3

US-10-073-293A-6 (1-111) x US-09-711-164-134 (1-333)

Qy 24 LeuArgLeuValGlyAsnAlaArgProThrLysArgGlyAlaValGlyLeuLeu 43
Db 299 GTGGCTTCAGGATCGCATCTGCAGACCTTTTCACGCGGCATAATGCGCTTCATGCTC 240
Qy 44 -----AspThrIleGlyIleAlaSerIle 51
Db 239 GGGCCTTCGTCTACGAAATTTTCGTAACTTTTCAGATCGTCAATGTCCAGCGCCATCGTTG 180
Qy 52 CysAlaLeuLeuValSerThrAlaProGluValMethHisAspThrArgPheVal 71
Db 179 TGTTCAGCGTTA-----GCAATGGCAGATTCAGAACTTTCTTG 141
Qy 72 ProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrlsThrArgSerIle 91
Db 140 ACCAGTACAGCCGCTTTCTTGTGGTGTAGGTCAAAATATATCCAGAGCTCGCACACTTTC 81
Qy 92 IleProThrLeuSerAlaLeu-----AlaTyrlsGlyLeuAlaTrpLys 106
Db 80 TTACCGGATCAGGTTCAGTACGACACGCGGACCTTCTGACGAGNAGAACGAGCATGGCGA 21
Qy 107 ValMetAlaIleIle 111
Db 20 TGTTCAGCATAGT 6

RESULT 6
US-09-492-709A-227/c
; Sequence 227, Application US/09/492.709A
; Patent No. 6720139
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Alllyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001A
; CURRENT APPLICATION NUMBER: US/09/492.709A
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 227
; LENGTH: 333
; TYPE: DNA
; ORGANISM: E. Coli
US-09-492-709A-227

Alignment Scores:
Pred. No.: 0.317 Length: 333
Score: 77.00 Matches: 24
Percent Similarity: 41.9% Conservative: 20
Best Local Similarity: 22.9% Mismatches: 37
Query Match: 13.9% Indels: 24
DB: 3 Gaps: 3

US-10-073-293A-6 (1-111) x US-09-492-709A-227 (1-333)

Qy 24 LeuArgLeuValGlyAsnAlaArgProThrLysArgGlyAlaValGlyLeuLeu 43
Db 299 GTGGCTTCAGGATCGCATCTGCAGACCTTTTCACGCGGCATAATGCGCTTCATGCTC 240

```



```
Qy 44 -----AspThrIleGlyIleAlaSerIle 51
Db 239 GGGCTTCCTACGAAATTTTCGTAACCTTCAGATCGTCAATGTACGCCATCGTGG 180
Qy 52 CysAlaLeuLeuValSerThrAlaProGluValMetHisAspThrArgArgPheVal 71
Db 179 TGTTCAGCGTTA-----GCAATGGCAGATCCAGAACTTCTCTG 141
Qy 72 ProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleIle 91
Db 140 ACCAGTACAGCCGCTTCTCTGTTGGTGTAGTCAAAATATCCAGAGCGTCGACACTTTC 81
Qy 92 IleProThrLeuLeuSerAlaLeu-----AlaTyrGlyLeuAlaTyrLys 106
Db 80 TTACCGGAATCAGGTTCAGCAACAGCGCAACCTTCTGAGCAGGAAGACGAGCATGGCGA 21
Qy 107 ValMetAlaIleIle 111
Db 20 TGTTCAGCATAGT 6

RESULT 7
US-09-134-001C-2694
; Sequence 2694, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC-007
; FILE REFERENCE: GTC-007
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1997-11-08
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2694
; LENGTH: 1968
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2694

Alignment Scores:
Pred. No.: 5.23 Length: 1968
Score: 76.00 Matches: 34
Percent Similarity: 34.0% Conservative: 21
Best Local Similarity: 21.0% Mismatches: 39
Query Match: 13.8% Indels: 68
DB: 3 Gaps: 7

US-10-073-293A-6 (1-111) x US-09-134-001C-2694 (1-1968)
Qy 5 ValLeuLeuGlyLeuLeuValGlyVal----- 14
Db 1435 ATTATGTTACTTGGTCTGTGATGGAGCTATGATGGCAATGTGACATGGCGGTCCTATT 1494
Qy 15 -----AlaAsnTyrCysPheArgTyrLeuProLeuArgLeuArgValGlyAsnAlaArg 32
Db 1495 AATAAGCAGCTTATGATTT-----GCTACTGCTGCATTAACGGAAGGAATGCGACT 1548
Qy 33 ProThrLysArgGlyAlaValGlyIleLeuLeuAspThrIleGlyIleAlaSerIleCys 52
Db 1549 CCTATTACTGCGCGATGATGGTGGTATGATCTCTCCATTAGCTATCGCTACA----- 1602
Qy 53 AlaLeuLeuValVal----- 57
Db 1603 GCAATGTTAATCTTTAGBAGAAATTCATAAGACGCAAGAGGCTCAATTGTACCAAC 1662
Qy 58 -----SerThrAlaProGluValMetHisAspThr 67
Db 1663 TATGTCATGGGATTCATATTTTACTGAAGGTGCCATCCCATTCGCGAGCTGCCGATCCA 1722
Qy 68 ArgArgPheValProThr---LeuValGly----- 76
```

```
Db 1723 TTACGTGTGATTCCTCAATGATGTTGGCTCAGGTGGCTGGAGCAATTCGCTTTAGGT 1782
Qy 77 -----PheAlaValLeuGlyAlaSerPhe 84
Db 1783 TTAGGATCAAGTATCAAAAGCACCTCATGGTGGAAATTTTCGTCATTATTTGGTACTGATTTT 1842
Qy 85 -----TyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 1843 AACCATATATCTCAAACTCTTATGCTTTAGTTGGGTAGCTAGTTTCAGCTTTAATT 1902
Qy 101 TyrGly 102
Db 1903 TATGTT 1908

RESULT 8
US-09-710-279-3573/c
; Sequence 3573, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3573
; LENGTH: 4041
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3573

Alignment Scores:
Pred. No.: 14.4 Length: 4041
Score: 76.00 Matches: 34
Percent Similarity: 34.0% Conservative: 21
Best Local Similarity: 21.0% Mismatches: 39
Query Match: 13.8% Indels: 68
DB: 3 Gaps: 7

US-10-073-293A-6 (1-111) x US-09-710-279-3573 (1-4041)
Qy 5 ValLeuLeuGlyLeuLeuValGlyVal----- 14
Db 958 ATTATGTTACTTGGTCTGTGATGGAGCTATGATGGCAATGTGACATGGCGGTCCTATT 899
Qy 15 -----AlaAsnTyrCysPheArgTyrLeuProLeuArgLeuArgValGlyAsnAlaArg 32
Db 898 AATAAGCAGCTTATGATTT-----GCTACTGCTGCATTAACGGAAGGAATGCGACT 845
Qy 33 ProThrLysArgGlyAlaValGlyIleLeuLeuAspThrIleGlyIleAlaSerIleCys 52
Db 844 CCTATTACTGCGCGATGATGGTGGTATGATCTCTCCATTAGCTATCGCTACA----- 791
Qy 53 AlaLeuLeuValVal----- 57
Db 790 GCAATGTTAATCTTTAGAGAAATTCATAAGACGCAAGAGGCTCAATTGTACCAAC 731
Qy 58 -----SerThrAlaProGluValMetHisAspThr 67
Db 730 TATGTCATGGGATTCATTTTATTACTGAAGGTGCCATCCCATTCGCGAGCTGCCGATCCA 671
Qy 68 ArgArgPheValProThr---LeuValGly----- 76
Db 670 TTACGTGTGATTCCTTCAATGATGTTGGCTCAGGTGTGGCTCAGGTGTGGCTAGCTTTAGGT 611
Qy 77 -----PheAlaValLeuGlyAlaSerPhe 84
```



```
Db 610 TTAGGATCAAGTATCAAAAGCACCTCATGGTGGAAATTTTCGTCAATTTATTTGTTACTGATTTT 551
Qy 85 -----TyrLysThrArgSerIleIleProThrLeuLeuSerAlaLeuAla 100
Db 550 AACCATATATCTCAAACTCTTATTCCTTTAGTTGGGTAGCTTAGTTTTCAGCTTTTAATT 491
Qy 101 TyrGly 102
Db 490 TATGGT 485

RESULT 9
US-09-710-279-3989
; Sequence 3989, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3989
; LENGTH: 4342
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3989

Alignment Scores:
Pred. No.: 15.9 Length: 4342
Score: 76.00 Matches: 34
Percent Similarity: 34.0% Conservative: 21
Best Local Similarity: 21.0% Mismatches: 39
Query Match: 13.8% Indels: 68
DB: 3 Gaps: 7

US-10-073-293A-6 (1-111) x US-09-710-279-3989 (1-4342)

Qy 5 ValLeuLeuGlyLeuValGlyVal-----14
Db 172 ATTATGTTACTTGCTGTGTGATGGAGCTATGATGCAATTCACATGGCGGTCATTT 231
Qy 15 -----AlaAsnTyrCysPheArgTyrLeuProLeuArgLeuArgValGlyAsnAlaArg 32
Db 232 AATAAAGCAGCTTATGTTATTT-----GCTACTGCTGCATTAACGGAAGGAATGCAGCT 285
Qy 33 ProThrLysArgGlyAlaValGlyIleLeuLeuAspThrIleGlyIleAlaSerIleCys 52
Db 286 CCTATTACTGCAGCGATGATTTGGTGTGATTCCTCCATTAGCTATCGCTACA-----339
Qy 53 AlaLeuValVal-----57
Db 340 GCAATGTAATCTTTAGAGAATAATTCATAAGAGCAAGAGCTCAATTGTACCAAC 399
Qy 58 -----SerThrAlaProGluValMetHisAspThr 67
Db 400 TATGTCATGGATTCATTTATTTACTGAAGGTGCCATCCATTCGCAGCTGCCATCCA 459
Qy 68 ArgArgPheValProThr-----LeuValGly-----76
Db 460 TTACGTGTGATTCCTCAATGATGGTTGGTGTGAGTGTGGCTGAGCAATTCCTTTAGGT 519
Qy 77 -----PheAlaValLeuGlyAlaSerPhe 84
Db 520 TTAGGATCAAGTATCAAAAGCACCTCATGGTGGAAATTTTCGTCAATTTATTTGTTACTGATTT 579
Qy 85 -----TyrLysThrArgSerIleIleProThrLeuLeuSerAlaLeuAla 100
```

```
Db 580 AACCATATATCTCAAACTCTTATTCCTTTAGTTGGGTAGCTTAGTTTTCAGCTTTTAATT 639
Qy 101 TyrGly 102
Db 640 TATGGT 645

RESULT 10
US-09-270-767-14385/c
; Sequence 14385, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14385
; LENGTH: 3209
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14385

Alignment Scores:
Pred. No.: 14.1 Length: 3209
Score: 75.00 Matches: 35
Percent Similarity: 44.0% Conservative: 16
Best Local Similarity: 30.2% Mismatches: 33
Query Match: 13.6% Indels: 32
DB: 3 Gaps: 6

US-10-073-293A-6 (1-111) x US-09-270-767-14385 (1-3209)

Qy 5 ValLeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArgTyrLeuProLeu 24
Db 804 GTTCTGATAGTCTTCTCTCATCGTC-----AGAGGACTCGGTCTC 765
Qy 25 ArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGlyIleLeuLeuAsp 44
Db 765 CGATTTCAGATC-----CGAGTCCCAATCAATGGAGTCATCGGAATCGTCATC--- 718
Qy 45 ThrIleGlyIleAlaSerIleCys-----52
Db 717 ---ATCGGGGGCGCAGCTTTGCCGGCGGAGCTTCACAGCTTCGGAACCGCTTCTGT 661
Qy 53 -----AlaLeuValValSerThrAlaProGluValMetHisAspThrArg 68
Db 660 AGGCTTCACTCCGCGCTGCTCATCCAGCATCGCCCGCTCTCCAGGGCAAC 601
Qy 69 ArgPheValProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrLysThrArg 88
Db 600 CACCTCGTC-----CTCGCCCTCGCTCTCTCTGTCGGGTGCTTCACGGAACCGGA 550
Qy 89 SerIleIleIleProThrLeuLeuSerAlaLeuAlaTyrGlyLeuAla 104
Db 549 CAGATCGCTCTC---GAAGTCTTGTATGTACTTTCGTACCTTCTGGCG 505

RESULT 11
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
```

```
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-10-103-840A-2

Alignment Scores:
Pred. No.: 3.9e+05 Length: 4403765
Score: 74.50 Matches: 24
Percent Similarity: 48.3% Conservative: 19
Best Local Similarity: 27.0% Mismatches: 41
Query Match: 13.5% Indels: 6
DB: 3 Gaps: 1

US-10-073-293A-6 (1-111) x US-09-103-840A-2 (1-4403765)

Qy 23 ProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGlyLeu 42
Db 1250937 CCCTTGATGCTACTATTATGGCAACGCCAGACCGCCCACTGTGGCGCCACAGGGGGCGCCT 1250996

Qy 43 LeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAlaProGlu 62
Db 1250997 TGGTCAGCGGTCGCACTACTAGCTTGTCTGTGGCTTACCCTATGCCGAGA 1251056

Qy 63 ValMethHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeuGlyAla 82
Db 1251057 -----CAACGCCGGCTGCCGTGATGTGGTGGCGGTGGCGGTGGACAGCGCA 1251101

Qy 83 SerPheTyrlsThrArgSerIleIleProThrLeuLeuSerAlaLeuAlaTyrcly 102
Db 1251102 CCGGCCCAACGGCTGGTTCGACCGG-CTCCCCCGCCTAACGCTACGGGTGCGCTTCGTC 1251160

Qy 103 LeuAlaTrpLysValMetAlaIle 111
Db 1251161 GTCTGCCAGGAGCTTTTCGGGTGATG 1251187

RESULT 12
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-05-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:
Pred. No.: 3.91e+05 Length: 4411529
Score: 74.50 Matches: 24
Percent Similarity: 48.3% Conservative: 19
Best Local Similarity: 27.0% Mismatches: 41
Query Match: 13.5% Indels: 6
DB: 3 Gaps: 1

US-10-073-293A-6 (1-111) x US-09-103-840A-1 (1-4411529)

Qy 23 ProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGlyLeu 42
Db 1251409 CCCTTGATGCTACTATTATGGCAACGCCAGACCGCCCACTGTGGCGCCACAGGGGGCGCCT 1251468

Qy 43 LeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAlaProGlu 62
Db 1251469 TGGTCAGCGGTCGCACTACTAGCTTGTCTGTGGCTTACCCTATGCCGAGA 1251528

Qy 63 ValMethHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeuGlyAla 82
Db 1251529 -----CAACGCCGGCTGCCGTGATGTGGTGGCGGTGGACAGCGCA 1251573

Qy 83 SerPheTyrlsThrArgSerIleIleProThrLeuLeuSerAlaLeuAlaTyrcly 102
Db 1251574 CCGGCCCAACGGCTGGTTCGACCGG-CTCCCCCGCCTAACGCTACGGGTGCGCTTCGTC 1251632

Qy 103 LeuAlaTrpLysValMetAlaIle 111
Db 1251633 GTCTGCCAGGAGCTTTTCGGGTGATG 1251659

RESULT 13
US-09-252-991A-6623/c
; Sequence 6623, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6623
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6623

Alignment Scores:
Pred. No.: 2.29 Length: 708
Score: 74.00 Matches: 40
Percent Similarity: 39.0% Conservative: 6
Best Local Similarity: 33.9% Mismatches: 41
Query Match: 13.4% Indels: 31
DB: 3 Gaps: 6

US-10-073-293A-6 (1-111) x US-09-252-991A-6623 (1-708)

Qy 20 ArgTyrlsProLeuArgLeuArg-ValGlyAsnAla----- 31
Db 568 CGATACCTTCTCGCGGTCTACCGCGCGCGCGCGGTACCGCGCGCGGTCTTCACCTT 509

Qy 32 -----ArgProThrLysArgGlyAlaValGlyLeuLe 43
Db 508 CGCGCGGTCTCTCGCGCGAGCTCGCGCGCGCGCGCGCGCGGTACCGCGCGGTCT 449

Qy 43 uAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSerThrAlaPro----- 61
Db 448 GTGCCTGCTCGCATCTCTCTCGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 389

Qy 62 -GluValMetHisAspThrArgArgPheValProThrLeuValGlyPhe-----AlaVa 79
Db 388 GGAACGCTCGCGCACCAACCGCGGTACCGCGCGCGGTACCGCGCGGTTCACCGCGGT 329

Qy 79 lLeuGly-----AlaSerPheTyrls-----ThrArgSerIleIlePr 93
Db 328 GGTCGGCCTGCTGCTGGCGCGGTGTACCGCGGTCTCTGACCGGTCTCTGCGGTCTCTGCGGCC 269
```


Percent Similarity: 44.4% Conservative: 18
 Best Local Similarity: 27.8% Mismatches: 38
 Query Match: 13.4% Indels: 22
 DB: 3 Gaps: 5

US-10-073-293A-6 (1-111) x US-09-603-208A-229 (1-729)

```

QY      5 ValLeuLeuLeuGlyLeuValGlyValAlaAlaSerThrCysPheArgTyrLeuProLeu 24
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     278 ATCATGCTCATCGGCTGGTATTGGTCTGGCCATGGATTACAGATCTTCCTCGTTACT 337
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     25 ArgLeuArg-----ValGlyAsnAlaArgProThr----- 34
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     338 CGTATGCGTGAGGGCTTCACCAAGGGCAAGACTGGGGCAACGCAACGTCGAATGGTTTC 397
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     35 LysArgGlyAlaValGlyIleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeu 54
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     398 AAGCACGGTGCC-----CGCGTGGTCACTGCTGCGCGCGCTG 433
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     55 LeuValValSerThrAlaProGluValMetHisAspThrArgArgPheValProThrLeu 74
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     434 ATCATGGTGTCTGTTCGGCGCATTCATAGCGGAGCATGGCGTTTATTAGACCATG 493
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     75 ValGlyPheAlaValLeuGlyAlaSerPheTyrIys-----ThrArgSerIleIle 91
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     494 ---GGCTTTGCTCTGGCGGTTGGTGTCTTCGATGCCTTCGTTGTTGGCATGATGATT 550
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     92 IleProThrLeuLeuSerAlaLeu 99
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     551 ATCCCTGCACAAATGTTCTCGCTT 574
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    
```

Search completed: February 16, 2006, 19:10:52
 Job time : 381.334 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:44:58 ; Search time 429.034 Seconds
(without alignments)
2139.461 Million cell updates/sec

Title: US-10-073-293A-6

Perfect score: 552

Sequence: 1 MSTEVLLGLLGVANVCFR.....IPTLLSALAYGLAWKNMII 111

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10073293/runat_15022006_120740_1106/app_query.fasta_1
-DB=Published Applications_NA_Main -QMT=fascap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes03p
-USER=US10073293 @CGN_1_1_2064 @runat_15022006_120740_1106 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA_Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	552	100.0	336	9	US-10-073-293A-5
2	98	17.8	172	6	US-10-360-201-18
3	94	17.0	670	7	US-10-424-599-80195
4	80.5	14.6	507	8	US-10-425-115-112500
5	79.5	14.4	3804	7	US-10-424-599-18810
6	78	14.1	1593	7	US-10-282-122A-27599
7	78	14.1	1879	3	US-09-728-260-1
c					Sequence 1, Appli

c	8	78	14.1	1879	3	US-09-798-412-1	Sequence 1, Appli
c	9	78	14.1	1879	5	US-10-325-917-1	Sequence 1, Appli
c	10	77.5	14.0	1326	7	US-10-282-122A-26760	Sequence 26760, A
	11	77.5	14.0	19988	7	US-10-672-787-10	Sequence 10, Appl
	12	77	13.9	330	6	US-10-287-274-57	Sequence 57, Appl
c	13	77	13.9	333	3	US-09-741-669-127	Sequence 127, App
c	14	77	13.9	333	3	US-09-912-020-227	Sequence 227, App
c	15	77	13.9	333	3	US-09-815-242-6257	Sequence 6257, Ap
c	16	77	13.9	333	6	US-10-287-274-134	Sequence 134, App
c	17	77	13.9	333	7	US-10-282-122A-1248	Sequence 1248, Ap
c	18	77	13.9	333	7	US-10-282-122A-6478	Sequence 6478, Ap
c	19	77	13.9	333	8	US-10-771-241-227	Sequence 227, App
c	20	77	13.9	654	7	US-10-437-963-70639	Sequence 70639, A
c	21	77	13.9	924	6	US-10-369-493-27216	Sequence 27216, A
c	22	77	13.9	1107	7	US-10-437-963-12803	Sequence 12803, A
c	23	77	13.9	10944	9	US-10-893-671-40	Sequence 40, Appl
c	24	76	13.8	330	7	US-10-282-122A-37364	Sequence 37364, A
c	25	76	13.8	333	3	US-09-815-242-9872	Sequence 9872, Ap
c	26	76	13.8	333	7	US-10-282-122A-7550	Sequence 7550, Ap
c	27	76	13.8	333	7	US-10-282-122A-39532	Sequence 39532, A
c	28	76	13.8	405	7	US-10-282-122A-37415	Sequence 37415, A
c	29	76	13.8	635	8	US-10-425-115-157416	Sequence 157416,
c	30	76	13.8	1544	7	US-10-092-900A-313	Sequence 313, App
c	31	76	13.8	1617	7	US-10-437-963-5392	Sequence 5392, Ap
c	32	76	13.8	1950	7	US-10-282-122A-34484	Sequence 34484, A
c	33	76	13.8	1968	7	US-10-724-972A-3483	Sequence 3483, Ap
c	34	75.5	13.7	826	8	US-10-425-115-106830	Sequence 106830,
c	35	75.5	13.7	999	9	US-10-501-282-5129	Sequence 5129, Ap
c	36	75.5	13.7	999	9	US-10-501-282-5131	Sequence 5131, Ap
c	37	75.5	13.7	999	9	US-10-501-282-5133	Sequence 5133, Ap
c	38	75.5	13.7	999	9	US-10-501-282-5135	Sequence 5135, Ap
c	39	75.5	13.7	999	9	US-10-501-282-5137	Sequence 5137, Ap
c	40	75.5	13.7	999	9	US-10-501-282-5139	Sequence 5139, Ap
c	41	75.5	13.7	999	9	US-10-501-282-5141	Sequence 5141, Ap
c	42	75.5	13.7	1754382	9	US-10-501-282-6651	Sequence 6651, Ap
c	43	75	13.6	3290	10	US-11-097-143-10172	Sequence 10172, A
c	44	75	13.6	5567	10	US-11-097-143-10171	Sequence 10171, A
c	45	74.5	13.5	1371	6	US-10-369-493-45062	Sequence 45062, A

ALIGNMENTS

RESULT 1

US-10-073-293A-5

; Sequence 5, Application US/10073293A
; Publication No. US20050239175A1
; GENERAL INFORMATION:
; APPLICANT: TABOLINA, EKATERINA
; APPLICANT: RYBAK, KONSTANTIN
; APPLICANT: KHOURGES, EVGENI
; APPLICANT: VOROSHILOVA, ELVIRA
; APPLICANT: GUSYATINER, MIKHAIL
; TITLE OF INVENTION: METHOD FOR PRODUCING L-AMINO ACID USING BACTERIA BELONGING TO THE
; FILE REFERENCE: 219594USO
; CURRENT APPLICATION NUMBER: US/10/073,293A
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: RU 2001103865
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: RU 2001104998
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: RU 2001104999
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: RU 2001117632
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: RU 2001117633
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Escherichia coli

```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(333)
; OTHER INFORMATION:
US-10-073-293A-5
Alignment Scores:
Pred. No.: 3,32e-68 Length: 336
Score: 552.00 Matches: 111
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-073-293A-6 (1-111) x US-10-073-293A-5 (1-336)
Qy 1 MetSerTyrGluValLeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArg 20
Db 1 ATGAGCTATGAGTTCTGCTGCTGGTTACTAGTTGGCGTGGGAATTAATGCTTCGCG 60
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 61 TATTGGCGCTGCGCTCGTGGGTAAATGCCGCCCAACCAACGTCGGCGGTAGGT 120
Qy 41 IleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValValSerThrAla 60
Db 121 ATTTTGTCTGCACACCAATTTGGCATCGCTCGATATGCGCTCTGTGGTTGTCTTACCGCA 180
Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 181 CCAGAAGTGATGCACGATACACGCGTTTCGTGCCACGCTGTCGGCTTCGCGGTACTG 240
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 241 GGTGCCAGTTTCTATPAAAACACGACGACATTAATCATCCCAACACTGCTTAGTGCCTGCC 300
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 301 TATGGGCTCGCTGGAAAGTGATGGCGATTATA 333

RESULT 2
US-10-360-201-18
; Sequence 18, Application US/10360201
; Publication No. US20030186292A1
; GENERAL INFORMATION:
; APPLICANT: MacNeil, Ian
; APPLICANT: Tiong, Choi Lai Yip
; APPLICANT: Brown, Kara
; TITLE OF INVENTION: METHODS FOR IDENTIFYING DNA MOLECULES THAT ENCODE A NATURAL PRODUCT
; TITLE OF INVENTION: HAVING BIOACTIVITY OR ENCODE A PROTEIN INVOLVED IN THE PRODUCTION
; TITLE OF INVENTION: PRODUCT HAVING BIOACTIVITY
; FILE REFERENCE: USAV2002/0007 USNP
; CURRENT APPLICATION NUMBER: US/10/360,201
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,083
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0213616.6
; PRIOR FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 172
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-360-201-18
Alignment Scores:
Pred. No.: 0.000283 Length: 172
Score: 98.00 Matches: 20
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 17.8% Indels: 0
DB: 6 Gaps: 0
```

```
US-10-073-293A-6 (1-111) x US-10-360-201-18 (1-172)
Qy 92 IleProThrLeuLeuSerAlaLeuAlaTyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 2 ATCCCAACACTGCTTAGTGCGCTATGGCTATGGGCTCGCTTGGAAAGTGATGGCGATTATA 61

RESULT 3
US-10-424-599-80195
; Sequence 80195, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 80195
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_43431C.1
US-10-424-599-80195
Alignment Scores:
Pred. No.: 0.00725 Length: 670
Score: 94.00 Matches: 37
Percent Similarity: 46.0% Conservatives: 27
Best Local Similarity: 26.6% Mismatches: 42
Query Match: 17.0% Indels: 34
DB: 7 Gaps: 5

US-10-073-293A-6 (1-111) x US-10-424-599-80195 (1-670)
Qy 2 SerTyrGluValLeuLeuLeuGlyLeuValGlyValAlaAsn---TyrCysPheArg 20
Db 210 TCTCTGGAGTTGCTATTGTCTACTCTTTTGTGTAAGTCACCGGATTTAGTGTCTCTTTGG 269
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 270 GAATGTCGGTGCATTCGAAACCTTTATGTGGGCAACCTATGTGTGGCGAGGATATAGGA 329
Qy 41 IleLeuLeuAspThrIleGlyIle----- 48
Db 330 AGTTTGGAACTACATTTGGTGGGATGTGACTGTGACATTTGTTGTCTCCGATAT 389
Qy 49 -----AlaSerIleCysAlaLeuValValSerThrAlaProGluVal 63
Db 390 CTCTGGTGTGGATATTCACTGATATAA-ATACTCATGCTGTTTAGTCAAGACCCCTGAAATT 448
Qy 64 MethHisAspThrArgArgPhe-----ValProThrLeuValGlyPheAlaVal 79
Db 449 TCTCATGCAGCTCGTGAGTATACGATATACCTTATCCAGCTCTATTGTTGGTCTGCTGTT 508
Qy 80 LeuGlyAla-----SerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSer 97
Db 509 CTTCAAGCTCTCACTCGCTACTTCCAGACTCAGATGATGATCTTCCCATGCTGTTTTCAGC 568
Qy 98 Ala-----LeuAlaTyrGlyLeuAlaTrpLysVal 107
Db 569 TCAATCACCGCACCTGTGTTTGCATGTTCTCTATTGTTGGGCTCTGGTTTAAAGTTG 625

RESULT 4
US-10-425-115-112500/c
; Sequence 112500, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
```



```

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27599
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Mycobacterium leprae
US-10-282-122A-27599

Alignment Scores:
Pred. No.: 4.62 Length: 1593
Score: 78.00 Matches: 31
Percent Similarity: 40.5% Conservative: 14
Best Local Similarity: 27.9% Mismatches: 32
Query Match: 14.1% Indels: 34
DB: 7 Gaps: 5

US-10-073-293A-6 (1-111) x US-10-282-122A-27599 (1-1593)
Qy 4 GluValLeu-----LeuLeuGlyLeuLeuValGlyValAlaAAsnTyrCysPheArg 20
Db 166 GAGGTGCTGTCGGGTGATCTGTTCGCTGATTGGCTGCTGCATCCACACACGGCGCTCGG 225
Qy 21 TyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLysArgGlyAlaValGly 40
Db 226 TATGCCGCGTGTGGCTC-----GAACGAGGTGCTGTGCGG 261
Qy 41 IleLeuLeuAspThrIleGlyLeuAlaSerIleCysAlaLeuLeuValValSerThrAla 60
Db 262 GTGCTTACCGACGTGCGAGGGGTCCCGAGCTGACAGCC-----CAGGCCACAGCGGTG 315
Qy 61 ProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValLeu 80
Db 316 CGGATACATACATACCGGAACCCCGC-----AGCGTGCTC 351
Qy 81 GlyAlaSerPheTyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
Db 352 GCGCGT-----CTTGCGCGCGCGGCA 372
Qy 101 TyrGlyLeuAlaTrpLysValMetAlaIleIle 111
Db 373 TATGACATCCCTCCACCGGATGACGGTGTGT 405

RESULT 7
US-09-728-260-1/c
; Sequence 1, Application US/09728260
; Publication No. US20020086980A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-319001
; CURRENT APPLICATION NUMBER: US/09/728,260
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/685,791
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/513,904
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/507,533
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/168,780
; PRIOR FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1879
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (113)....(1720)
US-09-728-412-1

Alignment Scores:
Pred. No.: 5.84 Length: 1879
Score: 78.00 Matches: 34
Percent Similarity: 46.4% Conservative: 17
Best Local Similarity: 30.9% Mismatches: 34
Query Match: 14.1% Indels: 25
DB: 3 Gaps: 5

US-10-073-293A-6 (1-111) x US-09-728-260-1 (1-1879)
Qy 2 SerTyrGluValLeuLeuLeuGlyLeuLeuVal-----GlyValAlaAAsnTyrCysPhe 19
Db 1850 AGCTCAGAGGTCAATTTATTAGGCTGCATTTTTTCTTAATACCTGTGTACACAGTAAGTGA 1791
Qy 20 ArgTyr-----LeuProLeuArgLeuArgValGlyAsnAla-ArgPr 33
Db 1790 GCGTACAGTAAAAAACCCCGCAGCATCCCTTCACATTCACACAGCGCTCAGCCTCGCG 1731
Qy 33 oThrLysArgGlyAlaValGlyIleLeuLeuAspThrIleGlyIleAlaSerIleCysAl 53
Db 1730 TGGTCTCTAGGAGCCCTCGT-----GTCCGTGTGTGCTGCTGCTGCTG 1686
Qy 53 aLeuLeuValVal-----SerThrAlaProGluValMetHisAspThrArgArgPheVa 71
Db 1685 TATTCCCGTATCCCTTCTCCCTGTCGCCAGCTCTCTTCGATCTTCGATCTTCGGAGGCCCGCT 1626
Qy 71 lProThrLeuValGlyPheAlaValLeuLeuGlyAlaSerPheTyrLysThrArgSerIleI 91
Db 1625 TCCTGCGGTAGTCTCGAAGCTCTCTTCGAGGCGCG----- 1589
Qy 91 elleProThrLeuLeuSerAlaLeuAla 100
Db 1588 ----CCGCTCCTCTCCGGGGGCTCGCT 1565

RESULT 8
US-09-798-412-1/c
; Sequence 1, Application US/09798412
; Publication No. US20030109428A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-327001
; CURRENT APPLICATION NUMBER: US/09/798,412
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 09/728,260
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/685,791
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/513,904
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/507,533
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/168,780
; PRIOR FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1879
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (113)....(1720)
US-09-798-412-1

Alignment Scores:
Pred. No.: 5.84 Length: 1879
Score: 78.00 Matches: 34

```


Percent Similarity: 46.4%
 Best Local Similarity: 30.9%
 Query Match: 14.1%
 DB: 3
 Indels: 25
 Gaps: 5

US-10-073-293A-6 (1-111) x US-09-798-412-1 (1-1879)

Qy 2 SerTyrGluValLeuLeuLeuGlyLeuVal-----GlyValAlaAsnTyrCysPhe 19
 Db 1850 AGCTCAGAGGTCATTTTATTAGGCTGCATTTTCTAATACTTGTACACAGGTATGTGA 1791
 Qy 20 ArgTyr-----LeuProLeuArgLeuValGlyAlaAsnAla-ArgPr 33
 Db 1790 GCGTACAGTAAAAAACCCGACGATCCCTTCACATTTACACAGCCTCAGCCTCGCG 1731
 Qy 33 oThrLysArgGlyAlaValGlyLeuLeuAspThrIleGlyLeuAlaSerIleCysAl 53
 Db 1730 TGGTCTCTAGAGCCCTCGT-----GTCGGTGTTCGCTGCTGCTCG 1686
 Qy 53 aLeuLeuValVal-----SerThrAlaProGluValMetHisAspThrArgArgPheVa 71
 Db 1685 TATTCCCGTGATCCCTCTCTCCCTGTCGCCAGCTGTTCTGCATCTTCGGAGCGCCGCT 1626
 Qy 71 lProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleIl 91
 Db 1625 TCCTGCGGTAGTTCCTGAAGCTCTCTTCGAGGCGCG-----1589
 Qy 91 elleProThrLeuLeuSerAlaLeuAla 100
 Db 1588 ----CCGCTCTCTTCGCGGGGCTCGCT 1565

RESULT 9

US-10-325-917-1/c
 ; Sequence 1, Application US/10325917
 ; Publication No. US20030113787A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; FILE REFERENCE: 07334-327001
 ; CURRENT APPLICATION NUMBER: US/10/325,917
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: US/09/798,412
 ; PRIOR FILING DATE: 2001-03-02
 ; PRIOR APPLICATION NUMBER: US 09/728,260
 ; PRIOR FILING DATE: 2000-12-01
 ; PRIOR APPLICATION NUMBER: US 09/685,791
 ; PRIOR FILING DATE: 2000-10-10
 ; PRIOR APPLICATION NUMBER: US 09/513,904
 ; PRIOR FILING DATE: 2000-02-25
 ; PRIOR APPLICATION NUMBER: US 09/507,533
 ; PRIOR FILING DATE: 2000-02-18
 ; PRIOR APPLICATION NUMBER: US 60/168,780
 ; PRIOR FILING DATE: 1999-12-03
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1879
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (113)...(1720)
 US-10-325-917-1

Alignment Scores:
 Pred. No.: 5.84 Length: 1879
 Score: 78.00 Matches: 34
 Percent Similarity: 46.4% Conservative: 17
 Best Local Similarity: 30.9% Mismatches: 34
 Query Match: 14.1% Indels: 25
 DB: 5 Gaps: 5

US-10-073-293A-6 (1-111) x US-10-325-917-1 (1-1879)
 Qy 2 SerTyrGluValLeuLeuLeuGlyLeuVal-----GlyValAlaAsnTyrCysPhe 19
 Db 1850 AGCTCAGAGGTCATTTTATTAGGCTGCATTTTCTAATACTTGTACACAGGTATGTGA 1791
 Qy 20 ArgTyr-----LeuProLeuArgLeuValGlyAlaAsnAla-ArgPr 33
 Db 1790 GCGTACAGTAAAAAACCCGACGATCCCTTCACATTTACACAGCCTCAGCCTCGCG 1731
 Qy 33 oThrLysArgGlyAlaValGlyLeuLeuAspThrIleGlyLeuAlaSerIleCysAl 53
 Db 1730 TGGTCTCTAGAGCCCTCGT-----GTCGGTGTTCGCTGCTGCTGCG 1686
 Qy 53 aLeuLeuValVal-----SerThrAlaProGluValMetHisAspThrArgArgPheVa 71
 Db 1685 TATTCCCGTGATCCCTCTCTCCCTGTCGCCAGCTGTTCTGCATCTTCGGAGCGCCGCT 1626
 Qy 71 lProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleIl 91
 Db 1625 TCCTGCGGTAGTTCCTGAAGCTCTCTTCGAGGCGCG-----1589
 Qy 91 elleProThrLeuLeuSerAlaLeuAla 100
 Db 1588 ----CCGCTCTCTTCGCGGGGCTCGCT 1565

RESULT 10

US-10-282-122A-26760
 ; Sequence 26760, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,948
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 26760
 ; LENGTH: 1326
 ; TYPE: DNA

```

Qy      9 GlyLeuLeuValGlyValAlaAenTyrcysPheArgTyr----- 21
Db     2843 GGTGGCGCTCATCGCCTTAATACATGTTGCAGCGCTTTTACCACACTTTGCCAACATTGAT 2902
Qy     22 -----LeuProLeuArgLeuArgValGlcYasnAlaArgProThrLys 35
Db     2903 TTGNATTTAGATGCTGACATTACCCATGCAATTATTGGTGCAAATGTTTCATCCATGG--- 2959
Qy     36 ArgGlyAlaValGlyLeuLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeu 55
Db     2960 ----CTTGCCACTTTGATGAGTGTATTATCATCTTTTGCC-----ATG 2995
Qy     56 ValValSerThrAlaProGluValMetHis-----Asp 66
Db     2996 ATTTTAAATACCCGCATCAGTTATTATTAATCTCCTCGCCAGCGTTTTTCAGGTGATGAC 3055
Qy     67 ThrArgArgPheValProThrLeuValGlyPheAlaValLeuGly----- 81
Db     3056 AACCAGCGCTTAAATGATATTGAATGGCTTGGTTTGGTCGGCTTTATCTTTAAGCTTT 3115
Qy     82 AlaSerPheTyrlYsthrArgSerIlellelleProThr-----LeuLeu 96
Db     3116 GCAGGTTTTAAAAAACTGGTCTCAATCATGTTCCCATCATCTGGCTACATTTGGTATGTTG 3175
Qy     97 SerAlaLeuAlaTyrgLyLeuAlaTip 105
Db     3176 CTTCCTTGTGCTGTTGTTGGTCATGG 3202

RESULT 12
US-10-287-274-57
; Sequence 57, Application US/10287274
; Publication No. US20030181408A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSESIS
; FILE REFERENCE: ELITEA.008DV1
; CURRENT APPLICATION NUMBER: US/10/287,274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-287-274-57

Alignment Scores:
Pred. No.:          0.679           Length:          330
Score:              77.00           Matches:         24
Percent Similarity: 41.9%           Conservative:    20
Best Local Similarity: 22.9%        Mismatches:     37
Query Match:       13.9%            Indels:         24
DB:                6               Gaps:           3

US-10-073-293A-6 (1-111) x US-10-287-274-57 (1-330)

Qy      24 LeuArgLeuArgValGlyAsnAlaAargProThrLysArgGlyAlaValGlyIleLeuLeu 43
Db     22 GTGGCTTCAGATGCGATCTGCACGACCTTTTGGCAGCGCATAAATGCCCTCATGCTC 81
Qy     44 -----AspThrIleGlyIleAlaSerIle 51
Db     82 GGGCGCTTCGTACGAAAATTTTCGTAACTTTTCAGATCGTCAATGTCAGCGCCATCGTTG 141
Qy     52 CysAlaLeuLeuValSerThrAlaProGluValMetHisAspThrArgArgPheVal 71

```

Qy 72 ProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleIle 91
 Db 181 ACCAGTACAGCGCTTCTGTGTGTAGTCAAAATATCCAGAGCCTCGACACTTTC 240
 Qy 92 IleProThrLeuLeuSerAlaLeu-----AlaTyrGlyLeuAlaTrpLys 106
 Db 241 TTACCGCAATCAGGTTCAGCAACAGGCGAACCTTCTGAGCAGAGAACGAGCATGCGCA 300
 Qy 107 ValMetAlaIleIle 111
 Db 301 TGTTAGCGATAGTT 315

RESULT 13

US-09-741-669-127/c
 ; Sequence 127, Application US/09741669
 ; Patent No. US2002022718A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Forsyth, R. Allyn
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; TITLE OF INVENTION: Genes identified as required for
 ; TITLE OF INVENTION: proliferation of E. coli
 ; FILE REFERENCE: ELITRA.009A
 ; CURRENT APPLICATION NUMBER: US/09/741,669
 ; PRIOR FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: US 60/173005
 ; PRIOR FILING DATE: 1999-12-23
 ; NUMBER OF SEQ ID NOS: 481
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 127
 ; LENGTH: 333
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(333)
 US-09-741-669-127

Alignment Scores:

Pred. No.:	0.688	Length:	333
Score:	77.00	Matches:	24
Percent Similarity:	41.9%	Conservative:	20
Best Local Similarity:	22.9%	Mismatches:	37
Query Match:	13.9%	Indels:	24
DB:	3	Gaps:	3

US-10-073-293A-6 (1-111) x US-09-741-669-127 (1-333)

Qy 24 LeuArgLeuValGlyAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleLeuLeu 43
 Db 299 GTGCGCTTCAGGATCGCATCTGCAGACCTTTTGCAGCGGCATAATGCGCTTCATGCTC 240
 Qy 44 -----AspThrIleGlyIleAlaSerIle 51
 Db 239 GGGCCTTCGTCTACGAAATTTTCGTAACTTTTCAGATCGTCAATGTCAGCGCCATCGTTG 180
 Qy 52 CysAlaLeuLeuValSerThrAlaProGluValMethHisAspThrArgArgPheVal 71
 Db 179 TGTTACGGTTA-----GCAATGGCAGATTCAGAACTTTCTTG 141
 Qy 72 ProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleIle 91
 Db 140 ACCAGTACAGCGCTTCTGTGTGTAGTCAAAATATCCAGAGCCTCGACACTTTC 81
 Qy 92 IleProThrLeuLeuSerAlaLeu-----AlaTyrGlyLeuAlaTrpLys 106
 Db 80 TTACCGCAATCAGGTTCAGCAACAGGCGAACCTTCTGAGCAGAGAACGAGCATGCGCA 21
 Qy 107 ValMetAlaIleIle 111
 Db 20 TGTTAGCGATAGTT 6

RESULT 14

US-09-912-020-227/c
 ; Sequence 227, Application US/09912020
 ; Patent No. US20020045592A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Trawick, John
 ; APPLICANT: Forsyth, R. Allyn
 ; APPLICANT: Froelich, Jamie M.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
 ; FILE REFERENCE: ELITRA.001DV1
 ; CURRENT APPLICATION NUMBER: US/09/912,020
 ; CURRENT FILING DATE: 2001-07-23
 ; PRIOR FILING DATE: 09/492,709
 ; PRIOR APPLICATION NUMBER: 2000-01-27
 ; PRIOR FILING DATE: 1999-01-27
 ; NUMBER OF SEQ ID NOS: 485
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 227
 ; LENGTH: 333
 ; TYPE: DNA
 ; ORGANISM: E. Coli
 US-09-912-020-227

Alignment Scores:

Pred. No.:	0.688	Length:	333
Score:	77.00	Matches:	24
Percent Similarity:	41.9%	Conservative:	20
Best Local Similarity:	22.9%	Mismatches:	37
Query Match:	13.9%	Indels:	24
DB:	3	Gaps:	3

US-10-073-293A-6 (1-111) x US-09-912-020-227 (1-333)

Qy 24 LeuArgLeuValGlyAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleLeuLeu 43
 Db 299 GTGCGCTTCAGGATCGCATCTGCAGACCTTTTGCAGCGGCATAATGCGCTTCATGCTC 240
 Qy 44 -----AspThrIleGlyIleAlaSerIle 51
 Db 239 GGGCCTTCGTCTACGAAATTTTCGTAACTTTTCAGATCGTCAATGTCAGCGCCATCGTTG 180
 Qy 52 CysAlaLeuLeuValSerThrAlaProGluValMethHisAspThrArgArgPheVal 71
 Db 179 TGTTACGGTTA-----GCAATGGCAGATTCAGAACTTTCTTG 141
 Qy 72 ProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleIle 91
 Db 140 ACCAGTACAGCGCTTCTGTGTGTAGTCAAAATATCCAGAGCCTCGACACTTTC 81
 Qy 92 IleProThrLeuLeuSerAlaLeu-----AlaTyrGlyLeuAlaTrpLys 106
 Db 80 TTACCGCAATCAGGTTCAGCAACAGGCGAACCTTCTGAGCAGAGAACGAGCATGCGCA 21
 Qy 107 ValMetAlaIleIle 111
 Db 20 TGTTAGCGATAGTT 6

RESULT 15

US-09-815-242-6257/c
 ; Sequence 6257, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel

Fri Feb 17 09:07:55 2006

```

; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6257
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(333)
US-09-815-242-6257

```

```

Alignment Scores:
Pred. No.: 0.688 Length: 333
Score: 77.00 Matches: 24
Percent Similarity: 41.9% Conservative: 20
Best Local Similarity: 22.9% Mismatches: 37
Query Match: 13.9% Indels: 24
DB: 3 Gaps: 3

```

US-10-073-293A-6 (1-111) x US-09-815-242-6257 (1-333)

```

Qy 24 LeuArgLeuValGlyAsnAlaArgProThrLysArgGlyAlaValGlyLeuLeu 43
Db 299 GTGGCTTCAGGATCGATCTGCAGACCTTTTGCACGGCGCATAAATCGCTTCATGCTC 240
Qy 44 -----AspThrIleGlyLeuAlaSerIle 51
Db 239 GGGCTTCGTCACGAAATTTTCGTAACTTTTCAGATCGTCATGTCAGCGCCATCGTGTG 180
Qy 52 CysAlaLeuLeuValValSerThrAlaProGluValMetHisAspThrArgArgPheVal 71
Db 179 TGTTCAGCGTTA-----GCAATGGCAGATTCCAGAACTTTCTTG 141
Qy 72 ProThrLeuValGlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerIleIle 91
Db 140 ACCAGTACAGCGCTTTCTTGTGTGTAGGTCAAAATATCCAGAGCCTGCGACACTTTC 81
Qy 92 IleProThrLeuSerAlaLeu-----AlaTyrGlyLeuAlaTrpLys 106
Db 80 TTACCGCAATCAGGTACAGCAACAGGCGAACCTTCTGAGCAGAGAACGAGCATGGCGA 21
Qy 107 ValMetAlaIleIle 111
Db 20 TGTTCAGCGATAGTT 6

```

Search completed: February 16, 2006, 18:23:48
Job time : 435.034 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 16, 2006, 17:47:13 ; Search time 297.143 Seconds
(without alignments)
792.964 Million cell updates/sec

Title: US-10-073-293A-6

Perfect score: 552

Sequence: 1 MSVEVLGLLLGVANYCFR.....IPTLLSALAYGLAWKMAII 111

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7204252 seqs, 1061369211 residues

Total number of hits satisfying chosen parameters: 14408504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/abs/ABSSWEB spool/US10073293/runat 15022006 120743 1175/app_query.fasta 1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pt0 -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs03p
-USPR=US10073293 @CGN 1.1 579 @runat 15022006 120743 1175 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

- 1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:**
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:**
- 3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:**
- 4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:**
- 5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:**
- 6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq1:**
- 7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq1:**
- 8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq1:**
- 9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq1:**
- 10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:**
- 11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:**
- 12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq4:**
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	78	14.1	1400	12	US-11-136-527-6758
C 2	78	14.1	1879	12	US-11-136-527-2662
C 3	76	13.8	4041	8	US-10-793-626-3573
C 4	76	13.8	4342	8	US-10-793-626-3989

5	74	13.4	729	11	US-11-082-389-363	Sequence 363, App
6	72.5	13.1	641	6	US-09-925-065A-848859	Sequence 848859,
7	72.5	13.1	3920	8	US-10-821-234-254	Sequence 254, App
C 8	72.5	13.1	100000	12	US-11-124-367A-5084	Sequence 5084, Ap
9	71.5	13.0	340	6	US-09-925-065A-921304	Sequence 921304,
10	71.5	13.0	342	6	US-09-925-065A-921323	Sequence 921323,
C 11	71.5	13.0	343	6	US-09-925-065A-924427	Sequence 924427,
C 12	71.5	13.0	344	6	US-09-925-065A-922200	Sequence 922200,
C 13	71.5	13.0	632	6	US-09-925-065A-482686	Sequence 482686,
14	70	12.7	1632	12	US-11-052-554A-546	Sequence 546, App
15	69.5	12.6	653	6	US-09-925-065A-874878	Sequence 874878,
C 16	68	12.3	3994	12	US-11-136-527-210	Sequence 210, App
17	67.5	12.2	1673	12	US-11-136-527-2344	Sequence 2344, Ap
18	67	12.1	3105	8	US-10-392-234A-29	Sequence 29, Appl
19	66.5	12.0	532	6	US-09-925-065A-808746	Sequence 808746,
20	66.5	12.0	532	6	US-09-925-065A-859760	Sequence 859760,
21	66.5	12.0	534	6	US-10-793-626-2531	Sequence 2531, Ap
C 22	66.5	12.0	538	6	US-09-925-065A-778471	Sequence 778471,
23	66.5	12.0	695	8	US-10-453-372-889	Sequence 889, App
C 24	66.5	12.0	1593	8	US-10-995-805-3	Sequence 3, Appli
C 25	66.5	12.0	2401	6	US-09-925-065A-90971	Sequence 90971, A
C 26	66.5	12.0	3229	8	US-10-793-626-3867	Sequence 3867, Ap
C 27	66.5	12.0	37500	8	US-10-522-037-1	Sequence 1, Appli
C 28	66	12.0	639	6	US-09-925-065A-385624	Sequence 385624,
C 29	66	12.0	683	6	US-09-925-065A-678535	Sequence 678535,
C 30	66	12.0	683	6	US-09-925-065A-678536	Sequence 678536,
C 31	66	12.0	1269	12	US-11-098-686-9055	Sequence 9055, Ap
C 32	66	12.0	100000	12	US-11-124-367A-5024	Sequence 5024, Ap
C 33	66	12.0	394468	8	US-10-995-561-13473	Sequence 13473, A
C 34	66	12.0	1457619	12	US-11-098-686-8739	Sequence 8739, Ap
C 35	65.5	11.9	2550	12	US-11-136-527-2113	Sequence 2113, Ap
C 36	65	11.8	201	8	US-10-995-561-77301	Sequence 77301, A
C 37	65	11.8	1868	9	US-11-072-512-1745	Sequence 1745, Ap
C 38	65	11.8	98309	12	US-11-124-368A-2921	Sequence 2921, Ap
39	64.5	11.7	556	8	US-10-453-372-895	Sequence 895, App
40	64.5	11.7	593	8	US-10-453-372-893	Sequence 893, App
41	64.5	11.7	625	6	US-09-925-065A-805065	Sequence 805065,
C 42	64.5	11.7	630	6	US-09-925-065A-747940	Sequence 747940,
43	64.5	11.7	631	6	US-09-925-065A-805092	Sequence 805092,
44	64.5	11.7	655	8	US-10-453-372-891	Sequence 891, App
45	64.5	11.7	739	6	US-09-925-065A-879702	Sequence 879702,

ALIGNMENTS

RESULT 1
US-11-136-527-6758/c
; Sequence 6758, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6758
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-6758

Alignment Scores:	8.28	Length:	1400
Pred. No.:	78.00	Matches:	34
Score:	46.4%	Conservative:	17
Best Local Similarity:	30.9%	Mismatches:	34
Query Match:	14.1%	Indels:	25
DB:	12	Gaps:	5

Qy 101 TyrGly 102
 Db 490 TATGGT 485

RESULT 4

US-10-793-626-3989
 ; Sequence 3989, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: P03480US

; CURRENT APPLICATION NUMBER: US/10793.626

; CURRENT FILING DATE: 2004-03-04

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3989

; LENGTH: 4342

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: nucleic acid sequence

US-10-793-626-3989

Alignment Scores:

Pred. No.: 60.5 Length: 4342
 Score: 76.00 Matches: 34
 Percent Similarity: 34.0% Conservative: 21
 Best Local Similarity: 21.0% Mismatches: 39
 Query Match: 13.8% Indels: 68
 DB: 8 Gaps: 7

US-10-073-293A-6 (1-111) x US-10-793-626-3989 (1-4342)

Qy 5 ValLeuLeuLeuGlyLeuValGlyVal----- 14
 Db 172 ATTATGTTATCTGCTGTGATGGAGCTATGATGCAATTGACATGGGGGTCATT 231
 Qy 15 -----AlaAsnTyrCysPheArgTyrLeuProLeuArgLeuArgValGlyAsnAlaArg 32
 Db 232 AATAAAGCAGCTTATGATTT-----GCTACTGCTGATTAACGGAAGGAATGCAGCT 285
 Qy 33 ProThrLysArgGlyAlaValGlyLeuLeuAspThrLeuGlyLeuAlaSerIleCys 52
 Db 286 CCTATTACTGCAGCGATGATGGTGTATGATTCCTCCATTAGCTATCGCTACA----- 339
 Qy 53 AlaLeuLeuValVal----- 57
 Db 340 GCAATGTAATCTTTAGAGAAATTCACTAAGAGCAAAAGAGCTCAATTGTACCAAAAC 399
 Qy 58 -----SerThrAlaProGluValMetHisAspThr 67
 Db 400 TATGTCATGGATTCATTTATTTACTGAAGTGCATCCATTCGCGAGCTGCGATCCA 459
 Qy 68 ArgArgPheValProThr-----LeuValGly----- 76
 Db 460 TTACGTGTGATTCCTCAATGATGGTTGGCTCAGGTGGCTGAGCAATTGCTTTAGGT 519
 Qy 77 -----PheAlaValLeuGlyAlaSerPhe 84
 Db 520 TTAGGATCAAGTATCAAGACCTCATGTTGGGAATTTTCGTCAATTATGTAATGATTTT 579
 Qy 85 -----TyrLysThrArgSerIleIleIleProThrLeuLeuSerAlaLeuAla 100
 Db 580 AACCATATCTTCAAACTCTTATGCTTTAGTTGTTAGTGGTACGTTAGTTTCAGCTTAATT 639
 Qy 101 TyrGly 102
 Db 640 TATGGT 645

RESULT 5

US-11-082-389-363

; Sequence 363, Application US/11082389

; Publication No. US20050244935A1

; GENERAL INFORMATION:

; APPLICANT: Pompejus, Markus

; APPLICANT: Kroger, Burkhard

; APPLICANT: Schroder, Hartwig

; APPLICANT: Zelder, Oskar

; APPLICANT: Haberbauer, Gregor

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE TRANSPORT

; FILE REFERENCE: BGI-131CPCN

; CURRENT APPLICATION NUMBER: US/11/082.389

; CURRENT FILING DATE: 2005-03-16

; PRIOR APPLICATION NUMBER: US 09/603024

; PRIOR FILING DATE: 2000-06-23

; PRIOR APPLICATION NUMBER: US 60/141031

; PRIOR FILING DATE: 1999-06-25

; PRIOR APPLICATION NUMBER: US 60/143262

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: US 60/151281

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: DE 19930487.4

; PRIOR FILING DATE: 1999-07-01

; PRIOR APPLICATION NUMBER: DE 19930489.0

; PRIOR FILING DATE: 1999-07-01

; PRIOR APPLICATION NUMBER: DE 19931549.3

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19931550.7

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19932134.5

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19941379.7

; PRIOR FILING DATE: 1999-08-31

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 446

; SEQ ID NO 363

; LENGTH: 729

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (101)..(706)

; OTHER INFORMATION: RXS03042

US-11-082-389-363

Alignment Scores:

Pred. No.: 12.1 Length: 729
 Score: 74.00 Matches: 30
 Percent Similarity: 44.4% Conservative: 18
 Best Local Similarity: 27.8% Mismatches: 38
 Query Match: 13.4% Indels: 22
 DB: 11 Gaps: 5

US-10-073-293A-6 (1-111) x US-11-082-389-363 (1-729)

Qy 5 ValLeuLeuLeuGlyLeuValGlyValAlaAlaAsnTyrCysPheArgTyrLeuProLeu 24
 Db 278 ATCATGCTCATCGCCTGGTATTGTTGTCGCCCATGGATTACCATGATCTTCTCGTTACT 337
 Qy 25 ArgLeuArg-----ValGlyAsnAlaArgProThr----- 34
 Db 338 CGTATCGTCGAGGGCTTCACCAAGGCAAGACTGCGGCAACGCAACGTCGAATGGTTTC 397
 Qy 35 LysArgGlyAlaValGlyIleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeu 54
 Db 398 AACACGGTGGCC-----CGCGTGGTCACTGCTGGCGCGCTG 433
 Qy 55 LeuValValSerThrAlaProGluValMetHisAspThrArgArgPheValProThrLeu 74
 Db 555:|||||

```
Db 434 ATCATGGTGTCTGTGTTTCGGCGCATTCATAGCGCAGGACATGGCGTTTATTAAAGCAACATG 493
Qy 75 ValGlyPheAlaValLeuGlyAlaSerPheTyrIys-----ThrArgSerIleIle 91
Db 494 ---GCCTTTGCTCTGCGCGTGTCTGTGTTCTTCGATGCTCTTCGTGTGTTTCGCATGATGATT 550
Qy 92 IleProThrLeuLeuSerAlaLeu 99
Db 551 ATCCCTGCACAAATGTTCTCTGCTT 574

RESULT 6
US-09-925-065A-848859
; Sequence 848859, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 848859
; LENGTH: 641
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-848859

Alignment Scores:
Pred. No.: 16.2 Length: 641
Score: 72.50 Matches: 31
Percent Similarity: 40.0% Conservative: 13
Best Local Similarity: 28.2% Mismatches: 37
Query Match: 13.1% Indels: 29
DB: 6 Gaps: 6

US-10-073-293A-6 (1-111) x US-09-925-065A-848859 (1-641)
Qy 4 GluValLeuLeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArgTyrLeuPro 23
Db 213 GAGGTAAATGAAGCTGACCCCTTGTAGGGGCTGGC-----CCC 251
Qy 24 LeuArgLeu-----ArgValGlyAsn-----AlaArg 32
Db 252 CTGGGGGTCTTCTTGGGGTGTAGGCAGCACAGATGGAGGGGGCAGTGGCCAGA 311
Qy 33 ProThrIysArgGlyAlaValGlyLeuLeuAspThrIleGlyIleAlaSerIleCys 52
Db 312 GCACCTTCTGAAGGTAAATCTCGGGTGGCGATGACAGGGTGGGCTGAGGGGGCAGTCG 371
Qy 53 AlaLeuLeuValSer-----ThrAlaProGluValMetHisasp 66
Db 372 TCCCTAGTCTATGTCACCACCGAGTTTACCTCCCACTGCTTTGTCTTTGGAAGCAGAT 431
Qy 67 ThrArgArgPheValProThrLeuVal-----GlyPheAlaValLeuGlyAlaSerPhe 84
Db 432 TTGAAT-----CCTGCTCTTCCCTGGCAGCCAGACAGCCCTTGGCAAGCATTT 482
Qy 85 TyrIysThrArgSerIleIleIleProThr 94
Db 483 TACCTGAGACCCCTCCACTGTGCTGCTACC 512
```

```
RESULT 7
US-10-821-234-254
; Sequence 254, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 254
; LENGTH: 3920
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-254

Alignment Scores:
Pred. No.: 151 Length: 3920
Score: 72.50 Matches: 28
Percent Similarity: 52.9% Conservative: 17
Best Local Similarity: 32.9% Mismatches: 31
Query Match: 13.1% Indels: 9
DB: 8 Gaps: 7

US-10-073-293A-6 (1-111) x US-10-821-234-254 (1-3920)
Qy 2 SerTyrGluValLeuLeuLeuGlyVal---LeuLeuValGlyVal---AlaAsnTyrCysPhe 19
Db 588 TCGGTTGAAATGCTGATCCTCGGTCGCTGTGTTATTGGCCTCTTCTGCGGACTCTGCACA 647
Qy 20 ArgTyrLeuProLeuArgLeuArgValGlyAsnAlaArgProThrLys---ArgGlyAla 38
Db 648 GGTTTGTGGCCC-----ATGTACATTGGAGAGATCTGCGCTACTGCGCTCGGGGTGCC 701
Qy 39 ValGlyIleLeuLeuAspThrIleGlyIleAlaSerIleCysAlaLeuLeuValSer 58
Db 702 TTTGGC---ACTCTCAACAGCTGGGCATC-----GTTGTGGAATTTCTGTGGCCAG 752
Qy 59 Thr---AlaProGluValMetHisaspThrArgPheValProThrLeuValGlyPhe 77
Db 753 ATCTTTGGTCTGGAATTATCATCTTGGGTCTGAAGAGCTATNGCGCGCTGCTACTGGGTTT 812
Qy 78 AlaValLeuGlyAla 82
Db 813 ACCATCTTCTCTGCT 827

RESULT 8
US-11-124-367A-5084/c
; Sequence 5084, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0
```



```
; SEQ ID NO 5084
; LENGTH: 100000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-367A-5084

Alignment Scores:
Pred. No.:      8.17e+03      Length:      100000
Score:          72.50        Matches:      31
Percent Similarity: 42.5%    Conservative: 14
Best Local Similarity: 29.2% Mismatches:    33
Query Match:     13.1%      Indels:       28
DB:              12         Gaps:        2

US-10-073-293A-6 (1-111) x US-11-124-367A-5084 (1-100000)

Qy 1 MetSerTyrGluValLeuLeuGlyLeuValGlyValAlaAsnTyrCysPheArg 20
Db 78013 CTGCTTATGAACTGGGCTGCTTGGTCTCTTAGATTCTCAGCACCTCTTCAMTTTCA 77954
Qy 21 TyrLeuProLeuArgLeuArgValGlyValAsnAlaArgProThrLysArgGlyValAla-ValG1 40
Db 77953 GGGTGTGCTGCCCTCTTGTGCCACAGCTAGGAACCTCTCAAGACGAGCGAGTTGG 77894
Qy 40 YlleLeuLeuAspThrLleGlylleAlaSerlleCysAlaLeuValValSerThrAl 60
Db 77893 GGCAGCCATAGGCTAGCTAGCTCTCGTCT----- 77863
Qy 60 aProGluValMetHisAspThrArgArgPheValProThrLeuValGlyPheAlaValle 80
Db 77862 -----CCAGGAGTACGTCTTCATTGCTCATGTCTCAGTCCAGTGTCTT 77822
Qy 80 uGlyAlaSerPheTyrLysThrArgSerllelleilleProThrLeuLeuSerAlaLeuAl 100
Db 77821 AGGACTC-----TGGGTTTGTCTGTTTGT 77795
Qy 100 aTyrGlyLeuAlaTrp 105
Db 77794 TTTTGGTTTGTCTTGG 77779

RESULT 9
US-09-925-065A-921304
; Sequence 921304, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 921304
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-921304

Alignment Scores:
Pred. No.:      10          Length:      342
Score:          71.50      Matches:      17
Percent Similarity: 52.9%  Conservative: 20
Best Local Similarity: 24.3% Mismatches:    18
Query Match:     13.0%      Indels:       15
DB:              6         Gaps:        2

US-10-073-293A-6 (1-111) x US-09-925-065A-943123 (1-342)

Qy 51 IleCysAlaLeu-----LeuVal 56
Db 95 GTCTGTGCCCTCGAGGAGTTGCAAACTTAGTCTCGGGAATAAGACAATRCAGTGTAAAT 154
Qy 57 ValSerThrAlaProGluValMetHisAspThrArgArgPheValProThrLeu---Val 75
Db 155 ACGACTACAAGGAAGAAAGTGAATAATGGCACAAGCAGTGGTTATCGACCTTGACTGCA 214
Qy 76 GlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerllelleilleProThrLeu 95
Db 215 GGTGTGAATCACCTGGGAGCTTTTAAACACACAGTACCAGGTCTCTTCCAGAGGTT 274

Best Local Similarity: 24.3% Mismatches: 18
Query Match: 13.0% Indels: 15
DB: 6 Gaps: 2

US-10-073-293A-6 (1-111) x US-09-925-065A-921304 (1-340)

Qy 51 IleCysAlaLeu-----LeuVal 56
Db 93 GTCTGTGCCCTCGAGGAGTTGCAAACTTAGTCTCGGGAATAAGACAATRCAGTGTAAAT 152
Qy 57 ValSerThrAlaProGluValMetHisAspThrArgArgPheValProThrLeu---Val 75
Db 153 ACGACTACAAGGAAGAAAGTGAATAATGGCACAAGCAGTGGTTATCGACCTTGACTGCA 212
Qy 76 GlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerllelleilleProThrLeu 95
Db 213 GGTGTGAATCACCTGGGAGCTTTTAAACACACAGTACCAGGTCTCTTCCAGAGGTT 272
Qy 96 LeuSerAlaLeuAlaTyrGlyLeuAlaTrp 105
Db 273 CTGATTTTCATTGGTCTGGGTACAGCCTGG 302

RESULT 10
US-09-925-065A-943123
; Sequence 943123, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 943123
; LENGTH: 342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-943123

Alignment Scores:
Pred. No.:      10          Length:      342
Score:          71.50      Matches:      17
Percent Similarity: 52.9%  Conservative: 20
Best Local Similarity: 24.3% Mismatches:    18
Query Match:     13.0%      Indels:       15
DB:              6         Gaps:        2

US-10-073-293A-6 (1-111) x US-09-925-065A-943123 (1-342)

Qy 51 IleCysAlaLeu-----LeuVal 56
Db 95 GTCTGTGCCCTCGAGGAGTTGCAAACTTAGTCTCGGGAATAAGACAATRCAGTGTAAAT 154
Qy 57 ValSerThrAlaProGluValMetHisAspThrArgArgPheValProThrLeu---Val 75
Db 155 ACGACTACAAGGAAGAAAGTGAATAATGGCACAAGCAGTGGTTATCGACCTTGACTGCA 214
Qy 76 GlyPheAlaValLeuGlyAlaSerPheTyrLysThrArgSerllelleilleProThrLeu 95
Db 215 GGTGTGAATCACCTGGGAGCTTTTAAACACACAGTACCAGGTCTCTTCCAGAGGTT 274

Alignment Scores:
Pred. No.:      9.96       Length:      340
Score:          71.50      Matches:      17
Percent Similarity: 52.9%  Conservative: 20
```

```

Qy 96 LeuSerAlaLeuAlaTyrglyLeuAlaTrp 105
||| :||| :||| :||| :|||
Db 275 CTGATTCATTGGTCTGGGTACAGCCTGG 304

RESULT 11
US-09-925-065A-924427/c
; Sequence 924427, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 924427
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-924427

Alignment Scores:
Pred. No.: 10.1 Length: 343
Score: 71.50 Matches: 17
Percent Similarity: 52.9% Conservative: 20
Best Local Similarity: 24.3% Mismatches: 18
Query Match: 13.0% Indels: 15
DB: 6 Gaps: 2

US-10-073-293A-6 (1-111) x US-09-925-065A-924427 (1-343)
Qy 51 IleCysAlaLeu-----LeuVal 56
||| :||| :||| :||| :|||
Db 251 GTCTGTGCCCTCGAGGAGTGCAACTTAGCTCGGGAATAAGACAATRCAGTGTAAAT 192

Qy 57 ValSerThrAlaProGluValMetHisAspThrArgPheValProThrLeu---Val 75
||| :||| :||| :||| :|||
Db 191 ACGACTACAGGAAGAAAGTATTAATGGCAAAAGCAGTGGTTATCGACCTTGACTGCA 132

Qy 76 GlyPheAlaValLeuGlyAlaSerPheTyrlsThrArgSerllelleProThrLeu 95
||| :||| :||| :||| :|||
Db 131 GGTGGAAATCACCTGGGGAGCTTTAAAAAACACCACTACAGGCTCTTCTCCAGAGGTT 72

Qy 96 LeuSerAlaLeuAlaTyrglyLeuAlaTrp 105
||| :||| :||| :||| :|||
Db 71 CTGATTCATTGGTCTGGGTACAGCCTGG 42

RESULT 12
US-09-925-065A-922200/c
; Sequence 922200, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 922200
; LENGTH: 344
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-922200

Alignment Scores:
Pred. No.: 10.1 Length: 344
Score: 71.50 Matches: 17
Percent Similarity: 52.9% Conservative: 20
Best Local Similarity: 24.3% Mismatches: 18
Query Match: 13.0% Indels: 15
DB: 6 Gaps: 2

US-10-073-293A-6 (1-111) x US-09-925-065A-922200 (1-344)
Qy 51 IleCysAlaLeu-----LeuVal 56
||| :||| :||| :||| :|||
Db 251 GTCTGTGCCCTCGAGGAGTGCAACTTAGCTCGGGAATAAGACAATRCAGTGTAAAT 192

Qy 57 ValSerThrAlaProGluValMetHisAspThrArgPheValProThrLeu---Val 75
||| :||| :||| :||| :|||
Db 191 ACGACTACAGGAAGAAAGTATTAATGGCAAAAGCAGTGGTTATCGACCTTGACTGCA 132

Qy 76 GlyPheAlaValLeuGlyAlaSerPheTyrlsThrArgSerllelleProThrLeu 95
||| :||| :||| :||| :|||
Db 131 GGTGGAAATCACCTGGGGAGCTTTAAAAAACACCACTACAGGCTCTTCTCCAGAGGTT 72

Qy 96 LeuSerAlaLeuAlaTyrglyLeuAlaTrp 105
||| :||| :||| :||| :|||
Db 71 CTGATTCATTGGTCTGGGTACAGCCTGG 42

RESULT 13
US-09-925-065A-482686/c
; Sequence 482686, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482686
; LENGTH: 632
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-482686

Alignment Scores:
Pred. No.: 21.4 Length: 632
Score: 71.50 Matches: 17

```


This Page Blank (uspto)